The Dissertation Committee for Yu-Fu Chen certifies that this is the approved version of the following dissertation:

Electrophysiological Measurement of Temporal Integration in Listeners with Normal Hearing

Committee:

Craig A. Champlin, Supervisor

Chang Liu

Harvey M. Sussman

Tiffany A. Whittaker

Electrophysiological Measurement of Temporal Integration in Listeners with Normal Hearing

by

Yu-Fu Chen, B.S., Au.D.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2016

Dedication

This dissertation work is dedicated to my wife, Yen-Ju Chen, for your endless love, support, and encouragement during the challenges of graduate school in the University of Texas at Austin. I am truly blessed to have you in my life. This work is also dedicated to my father, Tsu-Chang Chen, my mother, Shu-Ling Huang, and my sister, Yu-Wen Chen. All I have and will accomplish are only possible because of your love and sacrifice.

Acknowledgements

I have been very fortunate to be part of a great community of graduate students in the University of Texas at Austin. This dissertation would not have been finished without the help of many individuals. Therefore, I wish to acknowledge my deepest appreciation to these people.

First and foremost, I would like to express my greatest gratitude to my mentor, Dr. Craig A. Champlin, throughout my years in the Au.D. and Ph.D. programs. As my dissertation supervisor, Dr. Champlin worked closely with me, giving me guidance and support throughout the study. Without your inspiration and encouragement, this dissertation would not have been accomplished.

I would also like to thank my dissertation committee members, Dr. Chang Liu, Dr. Harvey M. Sussman, and Dr. Tiffany A. Whittaker. Your questions and comments about my study were extremely beneficial in the improvement of the study and the completion of the manuscript.

To my brilliant wife, I would like to thank you for always being my first and best subject for the study and providing valuable feedback when I needed it. Without your help, the experiment would not have progressed so smoothly.

I am also truly grateful to my loving family in Taiwan. Even though we were not in the same country physically, through the phone you constantly spoke words of encouragement to me and kept me moving forward at all times.

Finally, special thanks go to my spiritual family in Austin. Especially Kelly and Patricia Brown, Dharhas and Jada Pothina, Christopher and Mariana Rau, and Abby

Rigdon, you were always there to cheer me up and stood by me through good times and bad. Your friendship makes my life in Austin a wonderful experience.

Electrophysiological Measurement of Temporal Integration in Listeners with Normal Hearing

Yu-Fu Chen, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Craig A. Champlin

Temporal integration refers to the phenomenon whereby the detection threshold of the stimulus decreases (improves) as the signal duration increases. The majority of studies of temporal integration have relied on behavioral methods. As a result temporal integration can be influenced by the subject's physical and psychological status and these factors may affect signal detection. In the present study, the measurement of the auditory steady-state response (ASSR) was used to investigate temporal integration in listeners with normal hearing. The stimuli were sinusoidally amplitude modulated (SAM) signals varying in the modulation frequency (40 Hz and 80 Hz) and duration (50 ms, 100 ms, 200 ms, 300 ms, 400 ms, and 800 ms). The carrier was 1-kHz tone, 4-kHz tone, or white noise. The ASSRs were analyzed across different stimulus conditions in terms of amplitude, phase, signal-to-noise ratio (SNR), and percentage of detected responses. The results showed that temporal integration was more clearly revealed when the ASSR was recorded with the 40-Hz modulation frequency as compared to 80-Hz modulation frequency. For 40-Hz modulation frequency, the amplitude of the ASSR increased over the first 200 milliseconds after the stimulus onset until reaching a steady-state plateau, and then dropped rapidly after the stimulus offset. Conversely, the phase strength (i.e., variability) decreased over the first 400 milliseconds and remained relatively stable after

that. For 80-Hz modulation frequency, the ASSR amplitude did not increase until approximately 200 milliseconds, beyond which the ASSR amplitude increased at the same rate as for 40-Hz modulation frequency. In addition, the ASSR phase was less stable across subjects, which suggests weaker responses overall. An exponential model fit the electrophysiological data best; however, a significant frequency effect on the time constant was not observed. These results suggest that both auditory midbrain and brainstem are able to integrate auditory information over the first 200 milliseconds of stimulus.

Table of Contents

List of Tables

List of Figures

I. INTRODUCTION

Human auditory system has evolved the capability to combine auditory information and extract cues rapidly over time. This ability to process relevant cues helps not only to perceive speech, but also to detect environmental sounds. For example, the detectability of the sound increases as the sound duration increases. It has been found that up to around 250 milliseconds, the longer the sound is, the softer the sound can be for a human ear to detect. However, once the sound duration is longer than 250 milliseconds, the detectability of the sound remains relatively constant. The phenomenon whereby the detection threshold of the stimulus decreases (improves) as the signal duration increases is known as temporal integration. Temporal integration is considered as one of the fundamental abilities of the human auditory system.

Studies have shown that temporal integration is influenced both by listener and stimulus factors. One example of a listener factor is the hearing sensitivity of the subject. A listener with hearing loss might find the detectability of a brief sound is more difficult than a listener with normal hearing. An example of a stimulus factor is the frequency of sound. It has been found that the detectability of a brief low-frequency sound is more difficult than a brief high-frequency sound. While the general characteristics of temporal integration are well known, relatively less is understood about the underlying mechanism.

Numerous theories based on behavioral results have been proposed to account for the phenomenon of temporal integration. Two representative models are the energy detection and multiple looks models. The energy detection model states that the detectability of a sound is based on the integrated power of the sound over a certain fixed interval. The multiple looks model says that samples of the signal are stored in memory and then combined to determine the sound detectability. Both models have strengths and

limitations. For example, the energy detection model is unable to explain the phenomenon that the detection threshold for two brief tones only followed the classical energy detection model when the inter-stimulus interval (ISI) was smaller than 5 milliseconds, beyond which the threshold would remain relatively constant (Viemeister & Wakefield, 1991). Additionally, Viemeister and Wakefield (1991) showed that the detection threshold for two brief tones with the ISI of 100 milliseconds would not be affected by the intervening noise. On the other hand, multiple looks model does not specify how each 'look' is defined or how it contributes to the signal detection. Buus (1999) showed that if each modulation period of a modulated stimulus corresponds to a "look", the model requires increasing weights for each look over the first 200 milliseconds after the stimulus onset in order to explain the actual slope of -10 dB/decade for the first 100 milliseconds in the integration function.

The majority of studies of temporal integration have relied on behavioral methods. A behavioral experiment requires the subject's behavioral response. In the present study, the measurement of the auditory steady-state response (ASSR) was used to investigate temporal integration instead. The ASSR is a type of physiological measurement. The advantage of examining temporal integration physiologically is that it does not require a listener's behavioral response and therefore objective measurement of temporal integration in the auditory neural system can be obtained.

The ASSR is the neural response that can be elicited by the modulated stimulus. A modulated stimulus is typically generated by multiplying two sinusoids of very different frequencies together. The duration of this modulated stimulus can be varied systematically. The higher carrier frequency determines the site of maximal stimulation in the cochlea and the lower modulation frequency determines the site of maximal responsiveness in the neural pathway. For example, a stimulus with 40-Hz modulation

frequency might generate an ASSR from the auditory midbrain (Hari et al., 1989; Herdman et al., 2002; Johnson et al., 1988; Kiren et al., 1994; Mäkelä et al., 1990; Spydell et al., 1985). On the other hand, a stimulus with 80-Hz modulation frequency might generate the ASSR in the brainstem region (Cone-Wesson et al., 2002; Herdman et al., 2002; Hari et al., 1989; Kiren et al., 1994; Mäkelä et al., 1990; Picton et al., 2003). In other words, using different modulation frequencies could potentially allow the researchers to investigate temporal integration in different levels of the auditory system.

In the present study, the ASSRs from subjects with normal hearing were elicited with sinusoidally amplitude-modulated (SAM) tones and white noises while varying the modulation frequency (40 Hz and 80 Hz) and the duration (50 ms, 100 ms, 200 ms, 300 ms, 400 ms, 800 ms) of the signal. For the SAM tones, the carrier frequency was either 1000 or 4000 Hz. The collected ASSRs were analyzed across different stimulus conditions in terms of amplitude, phase, signal-to-noise ratio (SNR), and percentage of detected responses. Consistent with behavioral data, it was hypothesized that the amplitude of the ASSR would increase as the stimulus duration increased. Additionally, it was predicted that ASSR would be larger when evoked by a low-frequency carrier with a low-frequency modulator. By measuring the effects of temporal integration on the ASSR objectively, we anticipate the results will help us understand the mechanisms underlying temporal integration at different sites within the auditory nervous system.

II. LITERATURE REVIEW

2.1. Temporal integration

2.1.1. DETECTING SINUSOIDS

Temporal integration refers to the phenomenon whereby the detection threshold of the stimulus decreases (improves) as the signal duration increases. For people with normal hearing, the auditory detection threshold decreases at an average rate of about 8 to 10 dB per decade increase in stimulus duration (or about 3 dB per doubling of stimulus duration) up to approximately 250 milliseconds, after which the threshold changes relatively little (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957). Moreover, temporal integration allows the listener to combine or integrate acoustic energy over time to improve the detectability of sounds. Because detectability is the first step in perceiving sounds, temporal integration is considered one of the fundamental abilities of audition.

In one of the first systematic studies of temporal integration, Hughes (1946) measured the detection threshold of tones with different durations (63 to 739 ms) and frequencies (from 250 to 4000 Hz in octave steps). He found that a linear model (discussed below) best described the relationship between the threshold intensity and tone duration. In the Hughes' model, I represents the threshold intensity of a tone pulse with duration t, I_∞ represents the threshold intensity of a tone pulse with infinite duration, and τ represents the time constant. The time constant indicates the time it takes to achieve a criterion value. In audition, it is the time needed to reach maximal hearing sensitivity or the lowest (best) hearing threshold. In effect the linear model says that as stimulus duration increases, the detection threshold decreases at a progressively slower rate.

$$\frac{I}{I_\infty} = 1 + \frac{\tau}{t}$$

2.1.2. ENERGY DETECTION MODELS

Over the years additional models of temporal integration based on the energy of the stimulus have been proposed that describe the time course of the integration. For example, Green, Birdsall, and Tanner (1957) advanced a power function model to explain the relation between the detection threshold and signal duration. They measured the detectability of 1000-Hz tones with different durations. When the signal intensity was held constant, Green and his colleagues found that there was an increase in the detectability of the tone as the signal duration was increased from 250 milliseconds to 3 seconds. In other words, the detection threshold of the tone decreased continuously as the signal duration increased. The power function proposed by Green et al. (1957) is shown here:

$$It^m = C$$

where *t* represents time, *I* represents the threshold intensity of a tone pulse with arbitrary duration, *C* is a constant, and *m* is the function exponent or slope when plotted in the log-log coordinates. The larger the function exponent, the greater the function slope is and the more rapid the improvement in threshold.

Two years later Plomp and Bouman (1959) proposed an exponential model of temporal integration. They measured behavioral detection thresholds of tone pulses with different frequencies (from 250 to 8000 Hz) and durations (from 0.5 ms to 10 sec). The stimulus was delivered to subjects with normal hearing. In general, the threshold decreased by 3 dB for every doubling of the tone pulse duration. The exponential model shown below best described the relationship between threshold intensity and tone pulse duration:

$$\frac{I}{I_\infty} = \frac{1}{1 - e^{\frac{-t}{\tau}}}$$

where I represents the threshold intensity of a tone pulse with duration t, I_∞ represents the threshold intensity of a tone pulse with infinite duration, and τ represents the time constant. Plomp and Bouman (1959) concluded that the exponential model fit the data better than the linear model proposed by Hughes (1946) or the power model advocated by Green et al. (1957).

Both the exponential and power function models suggest that detection threshold improves as the stimulus duration increases. For the exponential model, only the signal that is within the time frame indicated by the time constant will be integrated. On the other hand, for the power function model there is no end to the integration. All the models predict that once the stimulus duration is longer than several hundred milliseconds, the threshold improvement becomes very small and will approach to final value asymptotically (with slightly different rates).

Watson and Gengel (1969) measured tonal detection threshold as a function of the frequency from 125 to 8000 Hz in octave steps and duration from 16 to 1024 milliseconds in logarithmic steps. The stimulus was delivered monaurally with contralateral broadband noise at 30 dB SPL. Watson and Gengel (1969) confirmed that the exponential model proposed by Plomp and Bouman (1959) fit their data well.

In addition, Florentine, Fastl, and Buus (1988) examined temporal integration in listeners with normal hearing, cochlear hearing losses with different configurations, and simulated hearing losses. They used a behavioral procedure with feedback to measure the tonal detection threshold as a function of signal frequency (0.25, 1, 4, 14 kHz) and duration (from 2 to 500 ms in steps of a factor of 2). The results showed that the detection threshold decreased by 8 to 10 dB for every decade increase in signal duration for listeners with normal hearing and simulated hearing losses. However, Florentine et al.

(1988) did not find the horizontal segment of the integration function as predicted by the exponential model (Plomp & Bouman, 1959).

In order to unify the measurement methods of temporal integration and determine the most appropriate model to account for the data, Gerken, Bhat, and Hutchison-Clutter (1990) reviewed 20 studies and measured the behavioral detection threshold using a diverse set of stimuli. He compared the exponential model (Plomp & Bouman, 1959) to the power function model (Green et al., 1957) to determine which one provided a better fit to the data. For listeners with normal hearing, the slope of temporal integration function was less than -10 dB per decade increase in stimulus duration. In other words, the power function exponent was less than unity. The slopes for single- and multi-burst stimuli were not significantly different from each other. Increasing the inter-stimulus interval (ISI) had little effect on the detection threshold of the inter-burst interval stimulus, which is contradicted with the prediction from the exponential model proposed by Plomp and Bouman (1959), where the threshold should become poorer as the ISI increases. Alternatively, Gerken et al. (1990) concluded that the power function model best described the temporal integration and the horizontal segment of the integration function advanced by Plomp and Bouman (1959) needed to be re-evaluated.

Studies have shown that the time constant τ in the linear (Hughes, 1946) and exponential (e.g., Plomp & Bouman, 1959) models and exponent m in the power function (e.g., Green et al., 1957) model of temporal integration are affected by the stimulus frequency (Watson & Gengel, 1969; Gerken et al., 1990). As the stimulus frequency increases, the time constant decreases and the exponent of the function increases. As one example, Plomp and Bouman (1959) showed that the time constant varied from 0.38 seconds for 250-Hz tone to 0.15 seconds for 8000-Hz tone. Watson and Gengel (1969) also reported that the time constant varied from approximately 0.25 seconds at low

frequencies to 0.05 seconds at high frequencies. Further, Gerken et al. (1990) demonstrated that the power function exponent *m* decreased as the signal frequency decreased. However, other investigators have not found such relationship (Garner, 1947; Zwicker & Wright, 1963; Florentine et al., 1988). For example, Florentine et al. (1988) did not find a significant frequency effect on the slope of temporal integration function.

Even though those studies mentioned above generally agree that the threshold decreases as the stimulus duration increases, little else is known about the mechanism underlying temporal integration. In 1960, Green proposed energy detection model assuming that the detection of a signal is based on the integrated power of the signal over a fixed time period. Such an integration process is usually presumed to occur over several hundred milliseconds (Plomp & Bouman, 1959; Green et al., 1957). An energy integrator with a time constant can count for the phenomenon that the detection threshold decreases as the stimulus duration increases up to a few hundred milliseconds.

2.1.3. DETECTING MODULATED SIGNALS

Temporal integration has been investigated not only with the detection threshold of sinusoidal tones, but also with modulated signals. Sinusoidal tones have flat waveform envelopes. However, sinusoidal amplitude-modulated (SAM) signals have fluctuating waveform envelopes. Studies have shown that temporal integration can be observed with the SAM signals (Buus, 1999; Viemeister, 1979; Sheft & Yost, 1990). The modulation threshold refers to the smallest amplitude fluctuation a listener can detect. Researchers have reported that the detection and modulation thresholds of the SAM signal decrease as the signal duration increases (Buus 1999; Viemeister, 1979; Sheft & Yost, 1990).

Viemeister (1979) measured the modulation threshold for the SAM wideband noise as a function of modulation frequency and stimulus duration. He found that the

modulation threshold increased as the stimulus duration was decreased (from 1500 to 500 to 250 ms). The increased threshold occurred with the 250-ms stimulus at modulation frequencies below 60 Hz. Viemeister (1979) proposed the modulation detection model to account for the phenomenon of temporal resolution. In the modulation detection model, the auditory system is believed to contain a "leaky integrator" or low-pass filter. The time constant of the low pass filter is about 2.5 milliseconds and represents the temporal resolution of the auditory system. The smaller the time constant, the better the temporal resolution of the auditory system is.

In order to gain a more complete picture about how the modulation threshold changes with different stimulus durations, Sheft and Yost (1990) measured the modulation threshold for signals changing in durations from 1 to 256 modulation periods for people with normal hearing. The experimental results showed that in general the modulation threshold decreased as the signal duration increased, and the mean slope was about -7.5 dB for every decade increase in stimulus duration. In fact, Buus (1999) measured the detection threshold of the 20-Hz SAM 1000-Hz tone as a function of stimulus duration. He found that temporal integration functions for the SAM tone and steady tone did not differ significantly from each other for listeners with normal hearing. Therefore, the listeners likely use the same listening strategy for detecting both the modulated and steady tones.

Due to the fact that the time constant for temporal resolution estimated by the modulation detection model differs from the time constant for temporal integration estimated by the energy detection model by two orders of magnitude, de Boer (1985) designated this discrepancy as "resolution-integration" paradox. The resolution-integration paradox implies that the attempt to describe the auditory system as a single leaky integrator with one time constant is inappropriate. On the other hand, Green (1985)

did not consider the discrepancy between temporal integration and resolution as paradoxical. Instead, he treated the temporal integration and resolution as two ends of the temporal continuum. The auditory system can invoke different temporal processes depending on the demands of the task.

2.1.4. ALTERNATIVE TO THE ENERGY DETECTION MODELS

Although the energy detection model can account for most features of temporal integration, it cannot account for some of the phenomena observed in temporal integration. For example, Viemeister and Wakefield (1991) showed that the detection threshold for two brief tones only followed the classical energy detection model when the ISI was smaller than 5 milliseconds, beyond which the threshold would remain relatively constant. In addition, they also demonstrated that the detection threshold for two brief tones with the ISI of 100 milliseconds would not be affected by the intervening noise. These findings are contradicted with the prediction made by the exponential model (Plomp & Bouman, 1959) and power function model (Green et al., 1957). According to the exponential model (Plomp & Bouman, 1959), the detection thresholds for two brief tones should increase from -3 to -1.5 dB when the ISI increases from 25 milliseconds to 175 milliseconds, and reach 0 dB when the ISI is more than 500 milliseconds. Besides, the intervening noise that is within the integration window should affect the detection threshold. On the other hand, according to the power function model (Green et al., 1957), the detection threshold for two brief tones should remain at -3 dB for the ISI within 100 milliseconds, beyond which the threshold should increase to -1.5 dB.

In an attempt to explain the “resolution-integration” paradox, Viemeister and Wakefield (1991) proposed an alternative model in which temporal integration is resulted from the “multiple looks” by the subject instead of an integrator with a time constant. In

the multiple looks model, the auditory system samples the input stimulus at a high sampling rate (a short time constant) and stores these samples in the short-term memory. The listener can retrieve these samples selectively. When the signal duration increases, the number of samples increases and consequently the listeners' performance improves, too. The combination of a short time constant and selective processing allows the multiple looks model to explain the result from the pulsed tone pair experiment (Viemeister & Wakefield, 1991). However, even though multiple looks model are able to solve the resolution-integration paradox, the model does not specify how each 'look' is defined or how it contributes to the signal detection. If each modulation period of a modulated stimulus corresponds to a "look," Buus (1999) showed that the model requires increasing weights for each look over the first 200 milliseconds after the stimulus onset in order to explain the actual slope of -10 dB/decade for the first 100 milliseconds in the integration function.

Researchers have suggested that the site of temporal integration lies somewhere in the central auditory system (Zwislocki, 1960; Gerken et al., 1990). More specifically, Zwislocki (1960) postulated that temporal integration takes place in the higher levels of auditory system because the latencies of cochlear activities are relatively short compared to the time constants obtained from the behavioral studies. Besides, Gerken et al. (1990) also showed that the contribution of the signal rise/fall time to temporal integration was negligible compared to the plateau portion of the signal. Due to the fact that the lower levels of auditory system are usually more responsive to the rise time of the stimulus (Keidel et al., 1983), Gerken et al. (1990) hypothesized that the temporal integrator resides in the central auditory system. However, the possibility also exists that the temporal integrator is located in the peripheral auditory system. Evidence for this position comes from Florentine et al. (1988) who showed that temporal integration was affected

by a cochlear lesion. Nevertheless, possible interactions between the central and peripheral auditory systems through the efferent auditory pathway, such as the olivocochlear bundle (Warr, 1980, 1992; Sahley et al., 1997), cannot be ruled out completely as well.

All the models of temporal integration mentioned thus far have relied on behavioral methods for data acquisition. As a result temporal integration can be influenced by the subject's physical and psychological status and these factors may affect signal detection. For example, in the models proposed by Plomp and Bouman (1959) and Green et al. (1957), both describe that the detection threshold improves as the stimulus duration increases; however, they differ on the curvature where the final detection threshold is approached asymptotically, and these slight differences could be caused by the listeners' attention on the task. Therefore, if possible physical and psychological effects from the listeners can be factored out when the signal becomes extremely long, the true course of temporal integration might be revealed. A physiological measurement thus may provide new insights about temporal integration. The advantage of examining temporal integration physiologically is that it does not require a listener's behavioral response; therefore, objective measurement on temporal integration in the auditory neural system can be obtained.

2.2. Auditory evoked potentials

2.2.1. TRANSIENT RESPONSES

One way to measure temporal integration physiologically is to record an auditory evoked potential (AEP). The AEP refers to the electrical activities of the auditory neural pathway that occur in response to acoustic stimuli (e.g., Arnold, 2000). In humans the AEP is typically recorded through the surface electrodes placed on the scalp. When the

stimulus must be presented many times and the responses are averaged, a fluctuating waveform emerges and the on-going electrical noise declines through cancelation. The waveform of the AEP consist of a series of peaks and valleys, the amplitudes of which range from 0.1 to 10 μV and occur between 0.2 to 500 milliseconds after the onset of stimuli (Hood, 1998). The AEPs can be classified in several ways. One method of classifying the AEPs is based on the source of neural generators of the response. Examples of this method include the electrocochleography (ECochG), auditory brainstem response (ABR), and cortical auditory evoked potential (CAEP). Another way to classify the AEPs is according to the time of occurrence (latency). Examples of this method are the short-latency response (SLR), middle-latency response (MLR), and long-latency response (LLR).

2.2.2. AUDITORY STEADY-STATE RESPONSE (ASSR)

Another way to classify the AEPs is based on the type of response, such as transient and steady-state responses. The transient response refers to responses evoked by a brief stimulus (e.g., click or tone burst) presented at a low rate. This technique allows a response to subside before the next response is evoked. Alternatively, the steady-state response refers to the AEP evoked by a modulated stimulus presented at a relatively high rate. This approach does not allow the response to completely decay before the next stimulus arrives and a stable response develops. Regan (1989) suggested that the frequency components of the steady-state response have stable amplitudes and phases over a relatively long time period compared to those of transient responses.

The auditory steady-state response (ASSR) appears to be the superposition of multiple transient responses from neural activities that are phase-locked to successive stimuli (Bohorquez & Ozdamar, 2008; Galambos et al., 1981; Hari et al., 1989; Plourde

et al., 1991; Stapells et al., 1984), as illustrated in Figure 1. In Figure 1(a) each modulation period of the SAM stimulus evokes a transient response. Multiple transient responses, each delayed by a modulation period, are overlapped as shown in Figure 1(b). In Figure 1(c) the summation of each delayed response forms the ASSR eventually. However, there are also other studies opposing against this hypothesis through the ASSR simulation (Roß et al. 2002) and observation on recovery time course of the ASSR after anesthesia (Santarelli & Conti, 1999). In fact, Basar et al. (1987) proposed that the ASSR might be generated from neurons that are responsible for rhythmic activities. Another study done by Dimitrijevic and Roβ (2008) suggested that these oscillatory neurons might be located in the thalamo-cortical pathway.

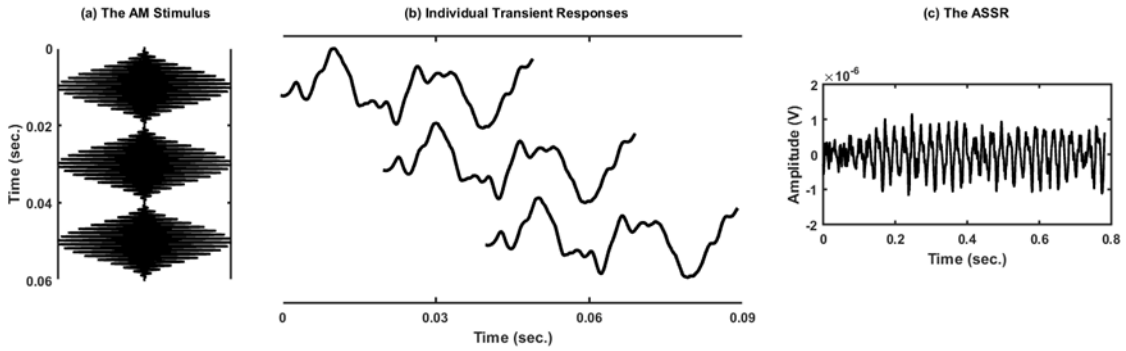

Figure 1. Shows SAM stimulus waveform in graph (a) and the transient neural response to each stimulus cycle in graph (b). The composite ASSR waveform in graph (c) represents the summed individual responses from graph (b).

The earliest studies of the ASSR can be traced back to 1960 (Geisler, 1960). However, it was not until Galambos and his associates (1981) recorded a series of AEPs with high stimulus rates that the ASSR gained broader attention. The stimulus used to evoke the ASSR is typically a modulated tone (additional details about the stimulus are provided on pages 16-18). The modulation is applied either to the amplitude or frequency

of the carrier, thus creating an amplitude- or frequency-modulated signal. The carrier frequency determines the general place of stimulation in the cochlea. In other words, it is the function of the neurons that are tuned to the carrier frequency. The modulation frequency of the modulated stimulus represents the rate at which neurons are synchronized to fire. In other words, if the neurons are responding to the modulated stimulus, the recorded ASSR would inherit the periodicity of phase-locked neural activities and the periodicity equals the modulation frequency of the modulated stimulus. When stimulated this way, the primary energy of a robust ASSR occurs at the modulation frequency.

The neural generators of the ASSR are mainly determined by the modulation frequency (Cone-Wesson et al., 2002; Hari et al., 1989; Herdman et al., 2002; Johnson et al., 1988; Kiren et al., 1994; Mäkelä et al., 1990; Picton et al., 2003; Spydell et al., 1985). If the modulation frequency exceeds 60 Hz, the responses are believed to originate from nuclei in the brainstem (Cone-Wesson et al., 2002; Hari et al., 1989; Herdman et al., 2002; Kiren et al., 1994; Mäkelä et al., 1990; Picton et al., 2003). However, if the modulation frequency is 20-60 Hz, the responses are hypothesized to come from neural activities in the auditory midbrain (Hari et al., 1989; Herdman et al., 2002; Johnson et al., 1988; Kiren et al., 1994; Mäkelä et al., 1990; Spydell et al., 1985). Finally, when the modulation frequency of the modulated stimulus is less than 20 Hz, neural generators of the ASSR are mainly located in the primary auditory cortex (Hari et al., 1989; Herdman et al., 2002; Mäkelä et al., 1990). In general, the ASSRs evoked by stimuli with higher modulation frequencies have shorter latency and are less susceptible to the maturation and arousal state of the subject due to that the primary sources of the ASSR locate in the brainstem (Lins & Picton, 1995; Cohen et al., 1991; Herdman et al., 2002; Levi et al., 1993), and as the fluctuation rate of the sound decreases, the latency of the ASSR

increases. By modulating the stimulus at different modulation frequencies, the ASSR allows the researchers to investigate different neural sites of the auditory system.

2.2.3. STIMULUS CONSIDERATIONS

A SAM signal used to evoke the ASSR is generated by multiplying a carrier signal by a second typically lower frequency modulation signal. The equation for a SAM pure tone is provided below:

$$f(t) = \sin(2 \times \pi \times CF \times t) \times [1 - m \times \cos(2 \times \pi \times MF \times t)] \text{ and}$$
$$m = \frac{Maximum\ AM\ amplitude - \ Minimum\ AM\ amplitude}{Maximum\ AM\ amplitude + \ Minimum\ AM\ amplitude}$$

In the equation, *t* represents the time, *CF* represents the frequency of the carrier tone, and *MF* represents the frequency of the modulation tone.

The extent the amplitude of the carrier signal is modulated by the modulation signal is described as the modulation depth. The method of calculating the modulation depth of the SAM signal is the same as the method to calculate the *m* value in the equation. For example, if a SAM signal has its amplitude fluctuating between 0 and 2, the modulation depth of the SAM is 100%, $[(2 - 0)/(2 + 0)] \times 100$. An example of a SAM tone and its amplitude spectrum is shown in Figure 2. For the SAM tone, the amplitude of the 1000-Hz carrier tone is modulated by the 40-Hz modulation tone, and the energy is primarily located at the carrier frequency, one sideband above the carrier frequency (carrier frequency + modulation frequency), and one sideband below the carrier frequency (carrier frequency – modulation frequency).

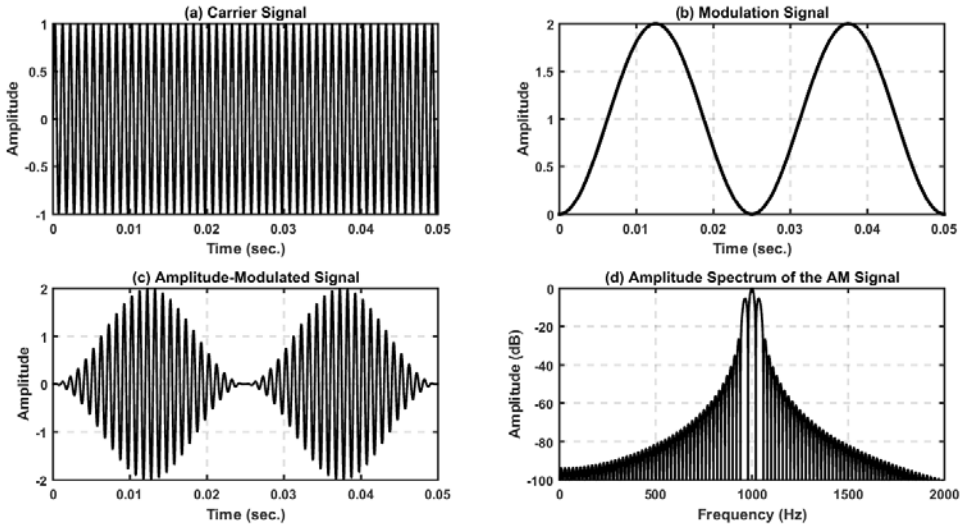

Figure 2. Example of the SAM tone and its amplitude spectrum. Graph (a) shows the 1000-Hz carrier tone, (b) the 40-Hz modulation tone, (c) a SAM tone composed of a 1000-Hz carrier tone modulated by a 40-Hz tone, and (d) the amplitude spectrum of the SAM tone.

However, the carrier signal of a SAM signal is not limited to the pure tone. In fact, the carrier signal can contain more than one frequency, such as the white noise. An example of a SAM white noise and its amplitude spectrum is shown in Figure 3. For the SAM white noise, the amplitude of the white-noise carrier is modulated by the 40-Hz modulation tone, and the energy is evenly distributed across all frequencies.

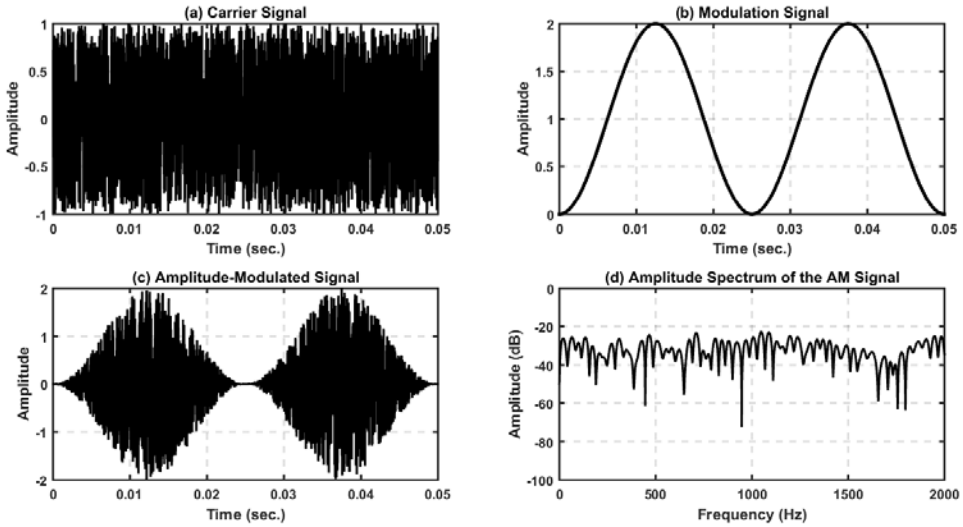

Figure 3. Example of the SAM white noise and its amplitude spectrum. Graph (a) shows the white-noise carrier, (b) the 40-Hz modulation tone, (c) a SAM noise composed of a white-noise carrier modulated by a 40-Hz tone, and (d) the amplitude spectrum of the SAM white noise.

2.2.4. RESPONSE DETECTION

The transient AEP is usually analyzed in the time domain and described in terms of the peaks and valleys in the waveform. However, the ASSR is most often analyzed in the frequency domain (Regan, 1989; John & Picton, 2000; Lins & Picton, 1995; Stapells et al., 1984). The AEP waveform is converted from the time domain to the frequency domain via the fast Fourier transform (FFT). The resulting amplitude and phase spectra are then available for analysis. Examples of analysis results from an ASSR are shown in Figure 4. In Figure 4, the amplitude at the modulation frequency represents the strength underlying the periodicity of phase-locked neural activities in the recorded ASSR. The higher the amplitude is at the modulation frequency compared to the amplitude at adjacent frequencies, the stronger the neural response is. The phase at the modulation frequency represents a relative timing when the neurons are synchronized to fire. If a neural response is present, the phase would be quite constant from one ASSR to the

other; otherwise, the phase would vary randomly. It has been suggested that the ASSR phase is more stable than the amplitude over the course of the recording (Champlin, 1992; Jerger et al., 1986). The frequency components other than the modulation frequency are usually considered as noise. The magnitude of the noise is commonly calculated as the average amplitude of the frequency components nearby the modulation frequency. The ratio of the amplitude of the neural response to that of the noise is usually defined as the signal-to-noise ratio (SNR) of the ASSR.

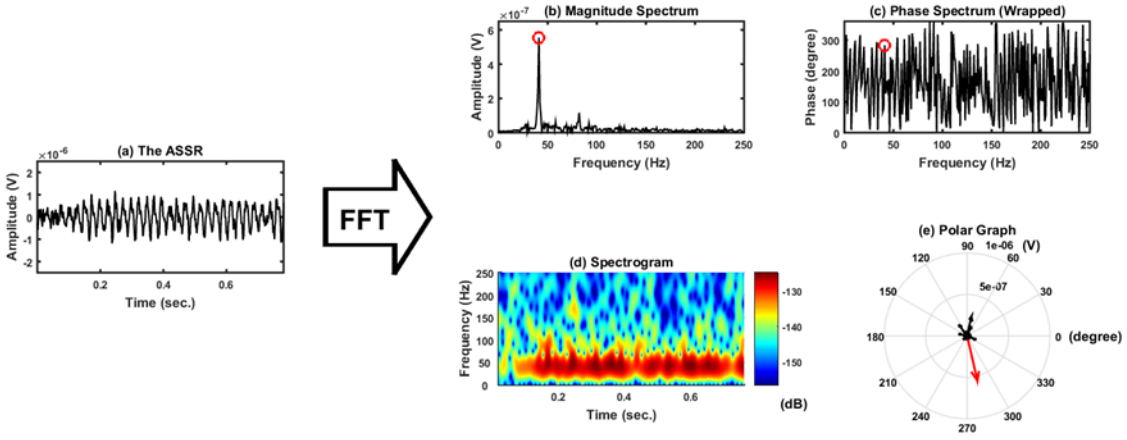

Figure 4. Analysis of the ASSR. The graph (a) shows a recorded time-domain ASSR to a SAM tone with 40-Hz modulation frequency, (b) the amplitude spectrum of the ASSR, with a red circle indicating the amplitude of the response at the modulation frequency, (c) the phase spectrum of the ASSR, with a red circle indicating the phase of the response at the modulation frequency, (d) the spectrogram of the ASSR, with the red color showing the primary energy distribution among different frequencies over time, and (e) the polar graph of the ASSR, with the length and angle of the red vector representing the amplitude and phase of the 40-Hz response respectively. The amplitude and phase of the 40-Hz response in the present example is 0.56 μV and 282 degrees, respectively.

The FFT yields information about the amplitude and phase of the response at discrete frequencies. Another useful transformation is the Hilbert transform (HT). The HT can be used to derive the instantaneous amplitude and phase of the ASSR over time (Roß et al., 2002). The effect of the HT on the ASSR is shown in Figure 5. Figure 5(a) shows a band-pass (30-50 Hz) filtered ASSR, whose instantaneous amplitude and phase over time are calculated through the HT. The instantaneous amplitude and phase of the ASSR as a function of the time is shown in Figure 5(b) and 5(c) respectively. On the other hand, the instantaneous frequency of the ASSR over time can also be calculated. After unwrapping the phase shown in Figure 5(c), the phase change between any two adjacent time points can be used to derive the slope, which represents the instantaneous angular frequency. The instantaneous frequency of the ASSR can be obtained after dividing the instantaneous angular frequency by 2π. The instantaneous frequency of the ASSR as a function of the time is shown in Figure 5(d). In the last, it is also possible to reconstruct an ASSR based on given instantaneous amplitude and phase over time. Both the FFT and HT are important signal processing tools. Researchers may choose one method to analyze the ASSR over the other depending on the purpose of the task.

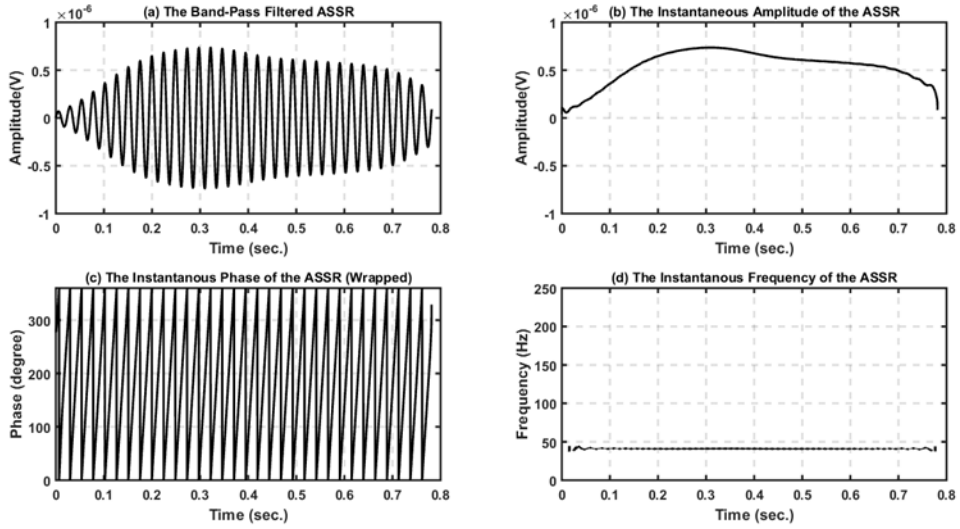

Figure 5 Analysis of the ASSR. Graph (a) shows a band-pass filtered ASSR to the SAM stimulus with 40-Hz modulation frequency. Graphs (b), (c), and (d) show the instantaneous amplitude, phase, and frequency of the ASSR over time, respectively.

2.3. Physiological measures of temporal integration

Several investigators have used neurophysiological methods to investigate the temporal integration. Onishi and Davis (1968) measured the amplitude and latency of the N1 response. The N1 wave is a component of the P1-N1-P2 response. The P1-N1-P2 response is considered as one type of the LLR and consists of a positive wave P1 at about 50 ms, a negative wave N2 at about 100 ms, and a positive wave P2 at about 200 ms following the stimulus onset. Onishi and Davis (1968) showed that in the condition of 30-ms rise time, the plateau length had no effects on the amplitude of the N1. In the condition of 3-ms rise time, the N1 amplitude increased as the plateau length increased up to 30 milliseconds and then remained relatively constant. In both conditions the latency of the N1 response decreased as the plateau length increased, especially at low intensity levels. When the plateau length was 2.5 seconds, a rise time shorter than 50

milliseconds had little effect on the N1 amplitude. Also, the amplitude and latency of the N1 evoked by the end of the signal were decreased compared to those evoked by the beginning of the signal. Onishi and Davis (1968) concluded that the N1 amplitude is determined by the first 30 milliseconds of the stimulus, and temporal integration might play a role in determining both the amplitude and latency of the N1.

Alain (1997) also measured the amplitude and latency of the N1 components (N1a, N1b, and N1c) and P2. The experimental results showed that the amplitudes of N1 components and P2 increased as the signal duration increased, and the signals with higher frequencies had shorter time constants. However, the different N1 components had different temporal integration functions, and components with longer latencies did not have longer time constants necessarily. On the other hand, the P2 latencies became longer as the signal duration increased. The scalp topography showed that the neural generators of the N1 components changed as a function of the signal duration and frequency; nevertheless, the neural generators of the P2 only changed as a function of the signal duration. Alain (1997) concluded that the N1 components and P2 might be generated by different types of neurons in the auditory cortex with integration times longer than 24 milliseconds.

Using a different approach, Sussman, Winkler, Ritter, Alho, and Näätänen (1999) measured the mismatch negativity (MMN). The MMN is considered as a type of the LLR and is usually obtained by randomly presenting perceptually different stimuli during a single AEP recording. One stimulus is presented more frequently (e.g., 80% of the presentations) and it is called the standard. The standard occurs sufficiently often that a pattern of expectation develops in the brain. The other stimulus is presented less often (e.g., 20% of the presentations) and it is called the deviant. The deviant disrupts expectation. Both standard and deviant stimuli can elicit the P1-N1-P2 response.

However, when the two stimuli are delivered randomly in a series of presentations, an additional negative peak known as MMN is appears between 100 and 300 milliseconds. In the Sussman et al. (1999) experiment, the MMN was measured as a function of the stimulus onset asynchrony (SOA) and the number of regularities. The SOA is the time period from the onset of the previous stimulus to that of the following stimulus. The number of regularities is the number of different standard stimuli that are used to establish expectation. The results showed that in the condition of one regularity, only one MMN was obtained if the SOA was 150 milliseconds; however, two MMNs were obtained if the SOA was 300 milliseconds. Sussman et al. (1999) explained that when the onsets of two deviant stimuli appeared successively in 150 milliseconds, the auditory system integrated them together; therefore, the regularity was only broken once and only one MMN was obtained. However, when the onsets of two deviant stimuli were separated by 300 milliseconds, the auditory system did not integrate them together anymore; therefore, the regularity was broken twice and two MMNs were obtained. On the other hand, in the condition of two regularities, two MMNs were still obtained even if the SOA was 150 milliseconds. This control condition demonstrated that two MMNs could still be generated even if the two deviant stimuli were 100-ms apart because two regularities were both broken. Sussman et al. (1999) concluded that the experimental results support the idea that temporal integration takes place within the window of about 200 milliseconds.

Loveless, Levänen, Jousmäki, Sams, and Hari (1996) recorded neuromagnetic evoked responses and measured the amplitude and latency of the N1 response as a function of the SOA of two tones. Their results showed that the response consisted of two N1s evoked by the two tones, and the second N1 consisted of two peaks. In general, the SOAs had little effects on the first N1 and early peak of the second N1; however,

significant effects of the SOAs were found on the late peak of the second N1, which were consistent with their previous investigations (Loveless & Hari, 1993; Loveless et al., 1989). Loveless et al. (1996) proposed that the second N1 might be comprised of an early ($N1^p$) and a late component ($N1^A$). The $N1^p$ was little affected by the SOAs. However, the amplitude and latency of the $N1^A$ increased when tones were presented at intervals less than about 250 milliseconds. Loveless et al. (1996) concluded that this enhancement effect might be resulted from a temporal integration process, which could store the auditory information holistically for about 200-300 milliseconds.

Gage and Roberts (2000) also used a neuromagnetic approach to record the N1 as a function of the stimulus type (constant intensity and constant energy) and duration. Similar results were found in both conditions of the stimulus type. The N1 amplitude increased as the stimulus duration increased up to 40 milliseconds and then remained relatively constant, but no significant effect of the stimulus duration was found on latency. Gage and Roberts (2000) concluded that the experimental results provide evidence that there is temporal integration of stimulus information within the window of about 40 milliseconds, and the amount of integration mainly depends on the stimulus duration instead of the stimulus intensity or energy.

Roβ, Picton, and Pantev (2002) measured the instantaneous amplitude and phase change of the neuromagnetic response over time to both 40-Hz SAM and non-modulated tone bursts. The experimental results showed that both types of stimuli generated the transient gamma-band response (GBR), N1, and sustained field (SF). However, the SAM tone burst generated an additional 40-Hz SSR. The GBR appears as successive periodic oscillations in the magnetic waveform and emerges between 20 and 200 milliseconds following the stimulus onset (Jacobson & Fitzgerald, 1997). The SF refers to the constant baseline shift during the recording of the neural response to the sound (Pantev et al.,

1994; Picton et al., 1978). In the Roß et al. (2002) experiment, the transient GBR was most obvious in the 100 milliseconds following both the onset and offset of the stimulus and its instantaneous amplitude and phase change over time varied from subject to subject. The transient GBR was shown to have close relationship with the MLR. On the other hand, the SSR was most obvious from 200 to 500 milliseconds and its instantaneous amplitude and phase change over time were very similar between subjects. The course of temporal integration could be observed on the SSR about 40 milliseconds after the stimulus onset, with the amplitude of the SSR increasing over 200 milliseconds until reaching the steady state, and then ends within 50-ms after the stimulus offset. The phase change of the SSR also gradually followed the modulation signal and became stable about 300 milliseconds after the stimulus onset. The magnetic source-localization procedures showed that the transient GBR and SSR were most likely generated at different locations in the primary auditory cortex. In addition, the real and simulated SSR were very similar in the first 70 milliseconds. However, the real SSR showed a drop in amplitude between 100 and 200 milliseconds after the stimulus onset, which was not observed in the simulated SSR. By manipulating the length of the ISI, The N1 amplitude decreased as the ISI decreased; nevertheless, the ISI had little effects on the 40-Hz SSR, transient GBR, and SF. Therefore, the decreases in amplitude between 100 and 200 milliseconds in the real SSR could not be explained by the inhibition from the N1 wave. Roß et al. (2002) concluded that the model proposed by Plomp and Bouman (1959) could explain the rising slope of the SSR better than the model proposed by Green et al. (1957), and the observed temporal integration on the SSR over the period of 200 milliseconds might take place in the primary auditory cortex. Roß and his associates suggested that the presence of the transient GBR before the SSR might explain the reason that smaller weights were given to the early phase of stimulus detection in the “multiple look” model,

They suggested the common assumption that the ASSR is the superposition of multiple transient responses needs to be re-evaluated due to the observed differences between the real and artificial responses. However, Roß et al. (2002) did not investigate how the temporal integration pattern observed in the auditory cortex differed from patterns acquired from different levels of auditory system, or whether there was a frequency effect on the course of temporal integration.

To sum up, several research studies have been published where temporal integration has been investigated by recording electric and magnetic responses. The physiological studies described above generally agree that the response amplitude increased as the stimulus duration increased, and that the signals with higher frequencies had shorter time constants. However, the latencies of different components did not show consistent trends. As the stimulus duration increased, the latency of some components increased, some decreased, and some remained unchanged. Moreover, response components with longer latencies did not have longer time constants necessarily.

It is worth noting that psychoacoustical and physiological results do not agree completely. For example, the time constants indicated by the physiological studies were about 30-40 milliseconds, as opposed to several hundred milliseconds of the time constant described in the behavioral measurement. In general, those AEPs elicited by transient stimuli do not permit the observation of temporal integration over an extended long time period (e.g., > 500 ms). Even though the neuromagnetic approach does allow for the observation of temporal integration for stimuli with longer durations, it does not permit the researchers to investigate different levels of auditory system. Therefore, the ASSR is a better candidate for studying temporal integration. The duration of the modulated signals used to obtain the ASSR can be varied systematically, and by varying the carrier frequency and the modulation frequency of the modulated signal, the effects of

the signal frequency on temporal integration at different levels of the auditory system can be investigated.

2.4. Specific aims & hypotheses

Temporal integration is a property of hearing that enables greater listener sensitivity. Studies of temporal integration indicate that signal detection improves as sound duration increases up to several hundred milliseconds after which there is little additional improvement. An index known as the time constant reflects the time it takes to reach maximal performance. Research has shown that the time constant decreases with increased sound frequency. This finding applies for both steady state and fluctuating sounds. Two models of temporal integration have been proposed. The 'energy detection' model states that the detection of a signal is based on the integrated power of the signal over a fixed time period. The 'multiple looks' model states that samples of the signal are stored in memory and then combined to determine the probability of signal detection. Both models have limitations. Studies have revealed that the time constant predicted by the energy detection model is longer than that typically found in time resolution studies. On the other hand, the multiple looks model does not specify how each 'look' is defined or how it contributes to the signal detection. Both models of temporal integration have relied on behavioral methods for their empirical data. As a result temporal integration can be influenced by the subject's physical and psychological status and these factors may affect signal detection. A physiological measurement such as the AEP avoids cognitive involvement and thus may provide new insights about temporal integration. However, the traditional AEPs elicited by transient stimuli do not permit the observation of temporal integration. Therefore, the ASSR is a better candidate for studying temporal integration because the duration of the modulated signals used to obtain the ASSR can be varied

systematically. Additionally, studies have shown that the phase of the ASSR is less influenced by the subject's arousal state than the amplitude; therefore, both the amplitude as well as the phase of the ASSR should be measured. It has also been suggested that time-varying sounds are processed in different parts of the auditory nervous system depending on the rate of fluctuation. Sound fluctuations are encoded at progressively higher (more rostral) sites as the rate decreases. Besides, as the fluctuation rate of the sound decreases, the latency of the ASSR increases. Consistent with behavioral results, we predict that the instantaneous amplitude of the ASSR will increase as the stimulus duration increases up to a point and then remain relatively constant. On the other hand, the instantaneous phase change rate of the ASSR will also gradually equal to that of the modulation signal as the stimulus duration increases up to a point. Also consistent with behavioral data, we predict that the time constant will decrease as the carrier frequency increases, and the white-noise carrier will have the smallest time constant. Finally, we predict that as the modulation frequency of the signal decreases, the time constant will increase. These hypotheses will be tested in conjunction with the following specific aims:

Aim 1: To assess temporal integration physiologically, the instantaneous amplitude and phase of the ASSR will be measured as a function of time.

Aim 2: To assess the effect of stimulus frequency, the instantaneous amplitude and phase of the ASSR will be measured by using different carrier frequencies.

Aim 3: To assess the general locus of the response, the instantaneous amplitude and phase of the ASSR will be measured by using different modulation frequencies.

By measuring the effects of temporal integration on the ASSR objectively, it is expected that the results will help us understand the mechanisms underlying temporal integration at different sites within the auditory nervous system.

III. METHODS

3.1. Subjects

Fifteen subjects (seven males and eight females; mean age = 23.87 years, SD = 4.55 years) were recruited for the study. None of the participants had any history of otological and neurological disorders. All the subjects passed the hearing screening in each ear at 10 dB HL at octave frequencies between 0.25 and 8 kHz. The test procedure was explained to each subject before informed consent was obtained. For each gender group, half were tested in the right ear and half were tested in the left ear. Subjects were paid for their participation.

3.2. Stimulus generation

The stimuli for the experimental conditions were SAM signals varying in the modulation frequency (40 Hz and 80 Hz) and duration (50 ms, 100 ms, 200 ms, 300 ms, 400 ms, and 800 ms). The carrier was 1000-Hz tone, 4000-Hz tone, or white noise. All the stimuli were modulated with 100% modulation depth. The stimuli were generated with a 16-bit digital-to-analog converter (Tucker-Davis Technologies, model QDA1) using a 50-kHz sampling rate. The stimuli were presented at a rate of 1.1864/s via an electrically shielded insert earphone (Etymotic, model ER-3A) at 80 dB SPL. For the control condition, no stimulus was delivered.

3.3. Acquisition

Two 3-hour experimental sessions were scheduled on different days for each subject. During the test participants were seated on a reclining chair in an acoustically and electrically shielded chamber while watching silent movies with subtitles through the course of the brainwave recording. The brainwaves were recorded using the disposable

electrode applied to the surface of the scalp. The skin at the electrode site was cleansed with a mild abrasive to ensure the inter-electrode impedances below 3 kΩ. The noninverting (active) electrode was placed on the forehead, the inverting (reference) electrode was placed on the earlobe of the test ear, and the ground electrode was placed on the earlobe of the non-test ear. For the experimental conditions, acquisition window was set at the length of the stimuli after the stimulus onset; for the control condition, acquisition window was set at 800 milliseconds from the start of the recording. Each brainwave was the average of 250 sweeps, and two brainwaves were collected for each condition. The test sequence for the 36 experimental conditions was pseudo-randomized. The experimental conditions were tested first, followed by the control condition that was conducted at end of the experiment. The brainwaves were differentially amplified (gain: 20) (RA4PA 4-Channels Medusa Preamp), filtered (passband: 30-300 Hz), digitized via a 16-bit analog-to-digital converter (Tucker-Davis Technologies, model QAD1) using a 25-kHz sampling rate, and then saved in the computer hard drive.

3.4. Analysis

3.4.1. FAST FOURIER TRANSFORM (FFT)

The digitized brainwaves were retrieved from the computer and analyzed offline. For each subject, two brainwaves of the same condition were averaged in the time domain. The FFT analyses of the average waveform were explained in different paragraphs for the experimental and control conditions.

For the experimental conditions, the averaged waveform for each condition was shaped with a Hamming window and converted into the frequency domain with the frequency resolution of 1 Hz through the FFT. The amplitude of the neural response was measured at the modulation frequency (i.e., 40 Hz or 80 Hz), and the amplitude of the

noise was estimated by averaging the amplitudes of the eight frequency components above and below the modulation frequency. The amplitudes of the neural response and noise were used to derive the SNR. The means and 95% confidence intervals of the response and noise amplitudes and the SNRs were calculated across subjects and compared across different conditions.

For the control condition, the 800-ms averaged waveform was first edited into 50-ms, 100-ms, 200-ms, 300-ms, 400-ms, and 800-ms waveforms from the onset. Each truncated waveform was then Hamming-windowed and converted into the frequency domain with the frequency resolution of 1 Hz through the FFT. The amplitudes of the frequency components at both modulation frequencies (i.e., 40 and 80 Hz) were measured, and the average amplitude of the 16 frequency components around each modulation frequency was calculated (they will be referred to as the "neural responses and noises" of the control condition). The amplitudes of the neural responses and noises were used to derive the SNRs. The means and 95% confidence intervals of the response and noise amplitudes and the SNRs were calculated across subjects and compared to the experimental conditions.

3.4.2. HILBERT TRANSFORM (HT)

In addition to the FFT, the averaged waveforms for the 800-ms conditions were band-pass filtered before performing the HT. A filter with the -3 dB bandwidth of 5 Hz centering at the modulation frequency was used before calculating the instantaneous amplitude of the waveform over time. The amplitude contour of the filtered waveform was scaled so that the mean amplitude in the interval between 300 and 700 milliseconds equals to the mean amplitude of all the amplitude contours in the same interval for the same stimulus condition. On the other hand, a filter with the -3 dB bandwidth of 20 Hz

centering at the modulation frequency was used before calculating the instantaneous phase of the waveform over time. The phase difference between the filtered waveform and the modulation signal was derived by subtracting the instantaneous phase of the waveform from that of the modulation signal. The phase contour was adjusted so that the mean phase difference in the interval between 300 and 700 milliseconds was equal to zero degree. The means and 95% confidence intervals of the amplitude and phase contours across subjects were calculated and compared across different conditions.

IV. RESULTS

4.1. Results from the pilot study

The results of the pilot study are briefly summarized here. Auditory steady-state responses (ASSRs) were acquired from 10 young adults using 40-Hz sinusoidally amplitude-modulated (SAM) 1000-Hz tones. The stimulus duration varied from 50 to 800 milliseconds in multiples of 2. Figure 6 shows the mean signal-to-noise ratio (SNR) as a function of stimulus duration with 95% confidence interval of the mean. In general, the mean SNR increased monotonically as the stimulus duration increased.

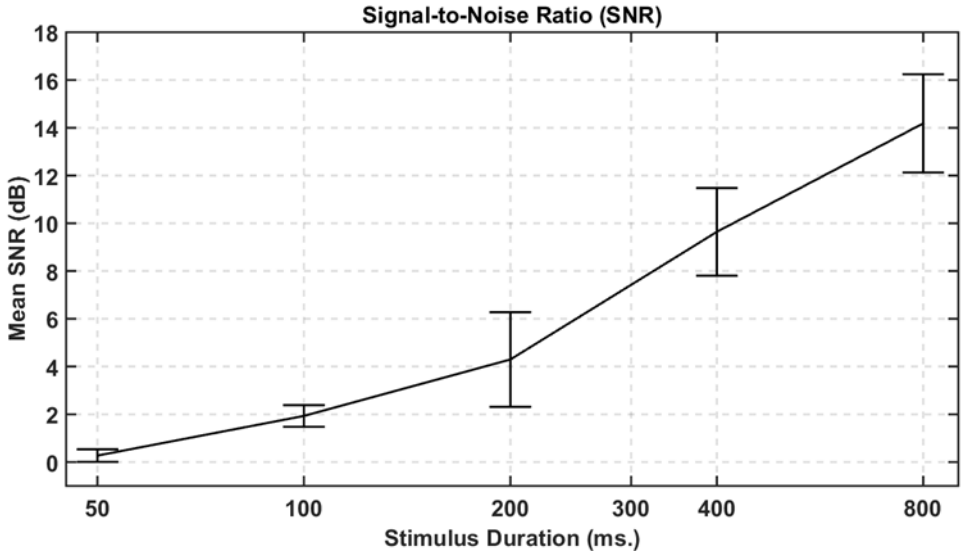

Figure 6 Mean SNR plotted as a function of stimulus duration on a logarithmic scale. Error bars indicate 95% confidence intervals of the mean.

The percentage of responses detected was calculated for all stimulus conditions. A response was considered as present if its amplitude was larger than the mean amplitude of the adjacent noise components plus two standard deviations of the mean. Figure 7 shows the percentage of detected responses as a function of stimulus duration. In general, the percentage of detected responses was 0% when the stimulus duration was shorter than

200 milliseconds. However, the percentage of detected responses exceeded 90% once the stimulus duration was longer than 200 milliseconds.

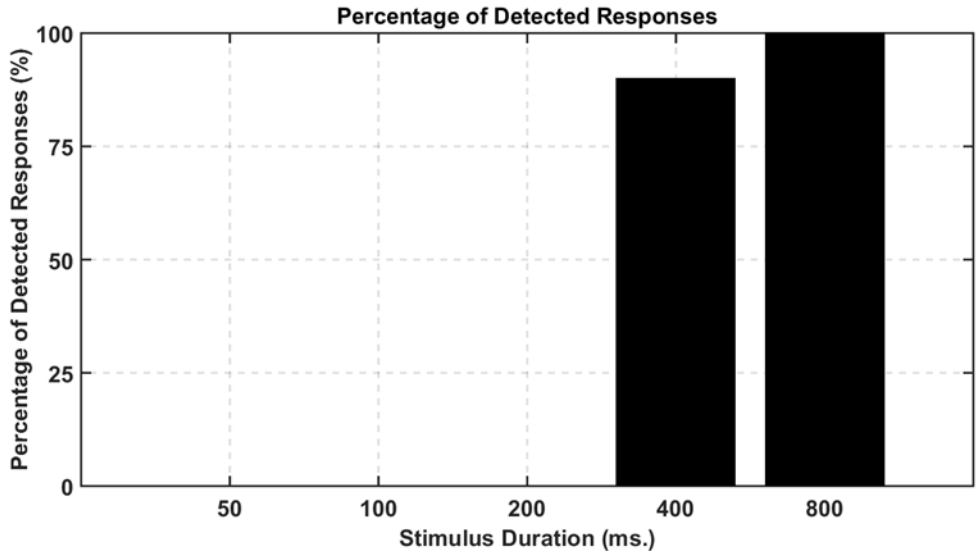

Figure 7 Percentage of detected responses plotted as a function of stimulus duration.

The individual 800-ms amplitude contours were then scaled so that the mean values between 300 and 700 milliseconds equal to the mean amplitude of all the amplitude contours in the same interval. Figure 8 shows the mean adjusted amplitude contour of the ASSR for the 800-ms stimulus condition. In general the amplitude of the ASSR increased gradually over the first 350 milliseconds after the stimulus onset until reaching a steady-state plateau, and then dropped rapidly after the stimulus offset within 50 milliseconds. The confidence interval of the amplitude contour during the first 350 milliseconds was wider compared to the last 450 milliseconds. The decrease in variability with increasing duration indicates a more robust neural response.

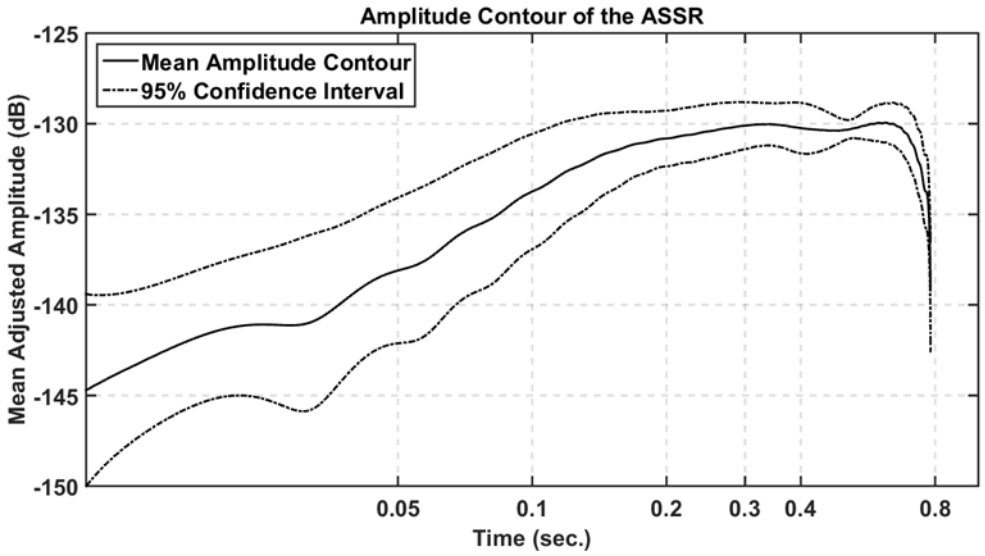

Figure 8 Mean instantaneous amplitude of the 800-ms ASSR plotted as a function of time on a logarithmic scale. Black line represents the mean amplitude contour of the ASSR with black dashed lines indicating 95% confidence interval of the mean.

Figure 9 shows the mean adjusted phase difference between the ASSR and 40-Hz modulation signal plotted as a function of time with 95% confidence interval of the mean for the 800-ms stimulus condition. Compared to the first 350 milliseconds, the confidence interval was narrower after 350 milliseconds, indicating stronger neural activity in response to the stimulus.

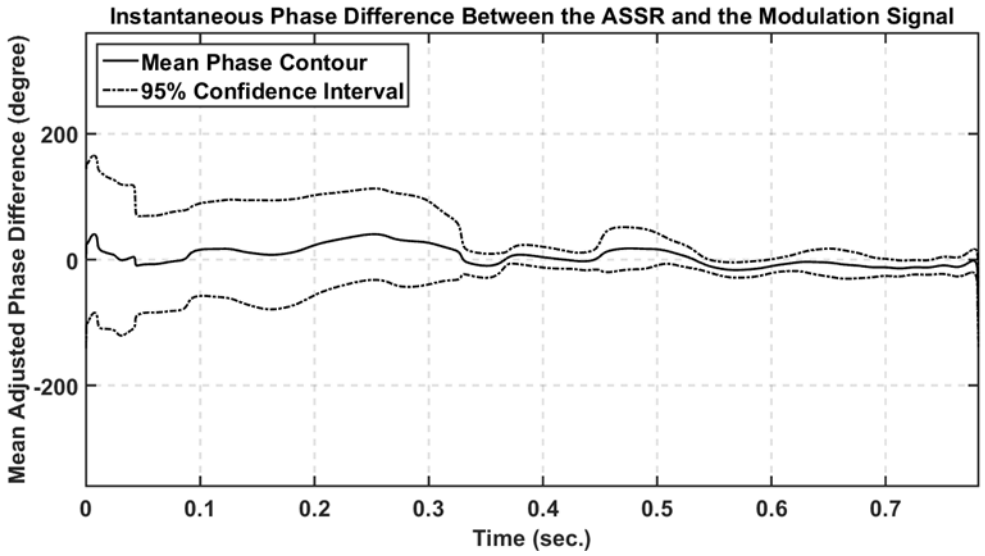

Figure 9 Mean phase difference between the 800-ms ASSR and 40-Hz modulation signal plotted as a function of time. Black line represents the mean phase contour of the ASSR with black dashed lines indicating 95% confidence interval of the mean.

To summarize the results of the pilot study, the SNR of the 40-Hz response increased as stimulus duration increased. More specifically, the ASSR amplitude increased monotonically over the first 350 milliseconds after the stimulus onset until reaching the steady state, and then dropped rapidly after the stimulus offset within 50 milliseconds. The confidence interval of the amplitude contour during the first 350 milliseconds was wider compared to the last 450 milliseconds. A similar result was observed for the phase contour. Taken together the amplitude and phase contours suggested the strength of the neural response grew stronger as the stimulus duration increased. In general, temporal integration could be observed on the ASSR over the first 350 milliseconds after the stimulus onset. These results are generally consistent with previous results from both psychoacoustic (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957) and physiologic (Roß et al., 2002) studies.

Although the pilot data showed that temporal integration could be observed in the ASSR, little else was revealed about the effects of stimulus frequency on temporal integration. Additionally, it was unknown whether temporal integration could be observed in different levels of auditory system. Therefore, the second (main) experiment was conducted. Besides to the stimulus conditions used in the pilot study, the main experiment also included stimuli with different carrier and modulation frequencies, duration of 300 milliseconds, and a control condition in which brainwave was recorded without playing any stimulus.

4.2. Results from the main experiment

The primary temporal factor in this study was stimulus duration. The effects of stimulus duration on the ASSR are considered from five different perspectives. In the first perspective, the focus is on the raw amplitude of the ASSR. In the second perspective, the ASSR amplitude is considered relative to the level of the background noise at adjacent frequencies. In the third perspective, the presence of a response is determined as the stimulus duration is varied. In the fourth and fifth perspectives, the instantaneous amplitude and phase of the ASSR were calculated as a function of duration. Each perspective is considered in the context of the two independent variables of carrier type and modulation frequency.

4.2.1. ASSR AMPLITUDE

The amplitude of the neural response at the modulation frequency was extracted through the fast Fourier transform (FFT) for both experimental and control conditions. The mean response amplitudes for the control condition were adjusted to zero and served as the baseline, and the mean amplitudes for the experimental conditions were adjusted by comparable amounts. Figure 10 shows the mean adjusted response amplitude plotted

as a function of stimulus duration with 95% confidence interval of the mean for all stimulus conditions. In general, the mean adjusted response amplitudes increased as the stimulus duration increased for all experimental conditions. The mean adjusted response amplitudes were larger for 1000-Hz carrier and 40-Hz modulation frequency. For modulation frequency of 40 Hz, the mean adjusted response amplitudes increased dramatically when stimulus durations were beyond 50 milliseconds; however, for modulation frequency of 80 Hz, the mean adjusted response amplitudes did not begin to increase until stimulus durations were longer than 200 milliseconds.

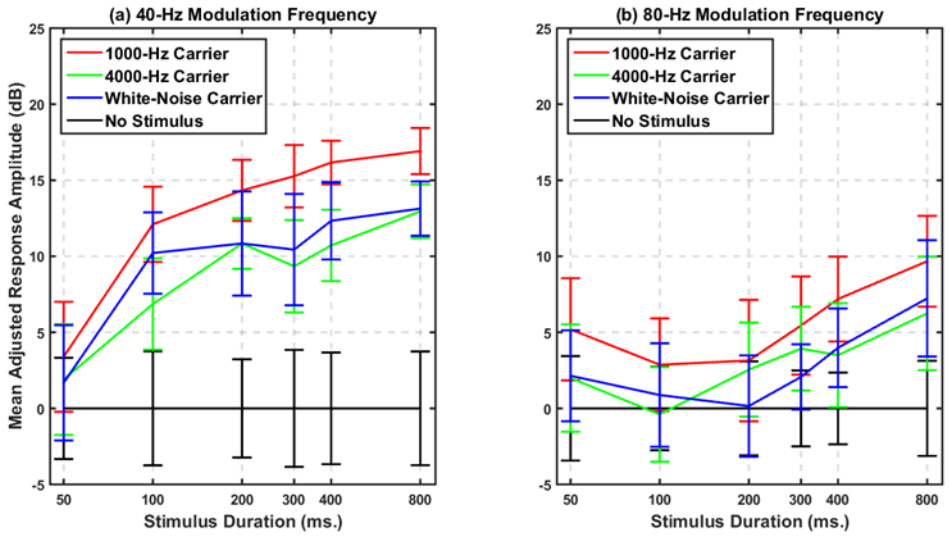

Figure 10 Mean adjusted response amplitude plotted as a function of stimulus duration on a logarithmic scale. Different panels represent different modulation frequencies (left: 40 Hz; right: 80 Hz). In each panel, different line colors represent different carrier types (red: 1000-Hz carrier; green: 4000-Hz carrier; blue: white-noise carrier; black: no stimulus). Error bars indicate 95% confidence intervals of the means.

A three-way repeated measures ANOVA [modulation frequency (2 levels) x carrier type (4 levels) x stimulus duration (6 levels)] with Geisser-Greenhouse corrected *p*-values was performed on the adjusted response amplitude. Modulation frequency,

carrier type, and stimulus duration were treated as categorical within-subject factors. The analysis showed significant main effects of modulation frequency [$F(1, 14) = 149.93$, $p < 0.05$, $\eta_p^2 = 0.91$], carrier type [$F(3, 42) = 44.448$, $p < 0.05$, $\eta_p^2 = 0.76$], and stimulus duration [$F(5, 70) = 19.887$, $p < 0.05$, $\eta_p^2 = 0.59$] on the adjusted response amplitude. Significant interactions were also found between modulation frequency and carrier type [$F(3, 42) = 14.297$, $p < 0.05$, $\eta_p^2 = 0.51$], between modulation frequency and stimulus duration [$F(5, 70) = 14.307$, $p < 0.05$, $\eta_p^2 = 0.51$], and between carrier type and stimulus duration [$F(15, 210) = 3.4974$, $p < 0.05$, $\eta_p^2 = 0.20$]. However, the interaction between modulation frequency, carrier type, and stimulus duration was not significant [$F(15, 210) = 1.6885$, $p > 0.05$, $\eta_p^2 = 0.11$]. The effects of modulation frequency and stimulus duration on the adjusted response amplitude were further explored by post hoc Dunn-Bonferroni analyses in different sessions for different experimental conditions.

Effects of carrier frequency

The pair-wise comparisons showed that for 1000-Hz carrier, the adjusted response amplitudes for 40-Hz modulation frequency were significantly greater than those for 80-Hz modulation frequency for stimulus durations longer than 50 milliseconds ($p < 0.05$). For 40-Hz modulation frequency, the adjusted response amplitudes for 1000-Hz carrier were significantly greater than those of the control conditions once stimulus durations exceeded 50 milliseconds ($p < 0.05$). Additionally, the adjusted response amplitudes for 1000-Hz carrier from 100 to 800 milliseconds were significantly larger than that at 50 milliseconds ($p < 0.05$), but there was no significant difference in adjusted response amplitudes between stimulus durations beyond 50 milliseconds ($p > 0.05$).

For 80-Hz modulation frequency, none of the adjusted response amplitudes for 1000-Hz carrier were significantly larger than those of the control conditions ($p > 0.05$).

Also, none of the adjusted response amplitudes for 1000-Hz carrier were significantly different between different stimulus durations either ($p > 0.05$).

Post hoc comparisons showed that for 4000-Hz carrier, the adjusted response amplitudes for 40-Hz modulation frequency were significantly higher than those for 80-Hz modulation frequency for stimulus durations of 200 and 800 milliseconds ($p < 0.05$). For 40-Hz modulation frequency, the adjusted response amplitudes for 4000-Hz carrier were significantly larger than those of the control conditions at stimulus durations of 200, 400, and 800 milliseconds ($p < 0.05$). Also, the adjusted response amplitude for 4000-Hz carrier was not significantly larger than the adjusted response amplitude at 50 milliseconds until stimulus duration was increased to 800 milliseconds ($p < 0.05$).

For 80-Hz modulation frequency, none of the adjusted response amplitudes for 4000-Hz carrier were significantly larger than those of the control conditions ($p > 0.05$). Also, none of the adjusted response amplitudes for 4000-Hz carrier were significantly different between different stimulus durations either ($p > 0.05$).

Post hoc comparisons showed that for white-noise carrier, the adjusted response amplitudes for 40-Hz modulation frequency were significantly higher than those for 80-Hz modulation frequency for all stimulus durations ($p < 0.05$) except for 50 and 800 milliseconds ($p > 0.05$). For 40-Hz modulation frequency, the adjusted response amplitudes for white-noise carrier were significantly larger than those of the control conditions at stimulus durations of 100, 400, and 800 milliseconds ($p < 0.05$). Also, the adjusted response amplitude for white-noise carrier was not significantly larger than that at 50 milliseconds until stimulus duration was increased to 800 milliseconds ($p < 0.05$).

For 80-Hz modulation frequency, none of the adjusted response amplitudes for white-noise carrier were significantly larger than those of the control conditions ($p >$

0.05). Also, none of the adjusted response amplitudes for white-noise carrier were significantly different between different stimulus durations either ($p > 0.05$).

For 40-Hz modulation frequency, there was no significant difference among the adjusted response amplitudes for 4000-Hz and white-noise carriers for all stimulus durations ($p > 0.05$), but the adjusted response amplitude for 1000-Hz carrier was significantly larger than those for 4000-Hz ($p < 0.05$) and white-noise ($p < 0.05$) carriers when the stimulus duration was 800 milliseconds. For 80-Hz modulation frequency, none of the adjusted response amplitudes were significantly different between different experimental conditions ($p > 0.05$).

In general, the adjusted response amplitudes for 40-Hz modulation frequency were larger than those for 80-Hz modulation frequency. For 40-Hz modulation frequency, the adjusted response amplitudes of the experimental conditions were larger than those of the control conditions once stimulus durations were longer than 50 milliseconds. The adjusted response amplitudes for 1000-Hz carrier were generally larger than those for 4000-Hz and white-noise carriers; however, not much difference in the adjusted response amplitudes between 4000-Hz and white-noise carriers was observed. For 80-Hz modulation frequency, the adjusted response amplitudes of the experimental conditions were not significantly different from those of the control conditions. None of the adjusted response amplitudes were significantly different between different experimental conditions either.

4.2.2. ASSR SIGNAL-TO-NOISE RATIO

The response amplitude was also considered relative to the amplitude of the background noise. The noise estimate was obtained by averaging the amplitudes of the frequency components immediately below and above the modulation frequency (i.e., 40

or 80 Hz). The amplitude at the modulation frequency was divided by the average noise amplitude to derive the SNR for each stimulus condition. The mean SNRs for the control condition were adjusted to zero and served as the baseline, and the mean SNRs for the experimental conditions were adjusted accordingly. Figure 11 shows the mean adjusted SNR plotted as a function of stimulus duration with 95% confidence interval of the mean for all stimulus conditions. In general, the mean adjusted SNRs increased as the stimulus duration increased for all experimental conditions. For modulation frequency of 40 Hz, the mean adjusted SNRs increased dramatically when stimulus durations were beyond 50 milliseconds; however, for modulation frequency of 80 Hz, the mean adjusted SNRs did not start to increase until stimulus durations were longer than 200 milliseconds.

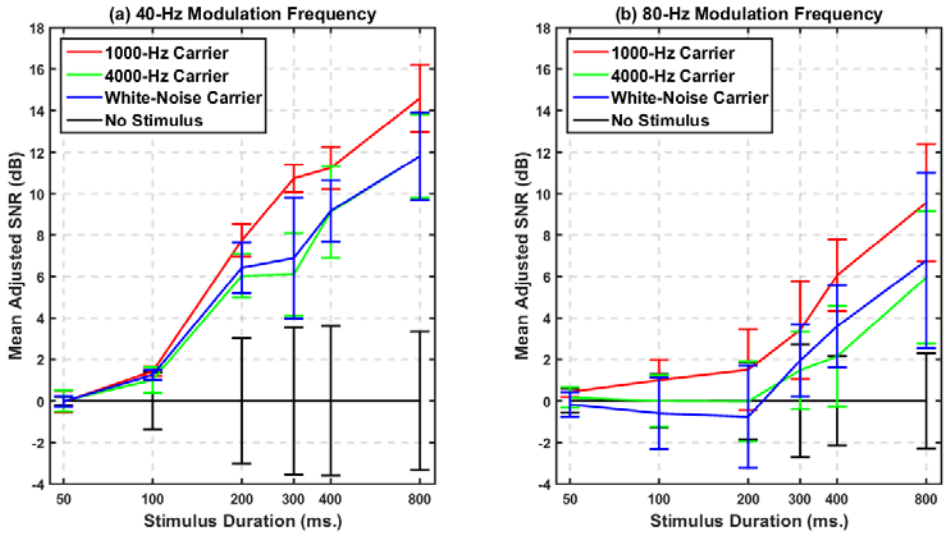

Figure 11 Mean adjusted SNR plotted as a function of stimulus duration on a logarithmic scale. Different panels represent different modulation frequencies (left: 40 Hz; right: 80 Hz). In each panel, different line colors represent different carrier types (red: 1000-Hz carrier; green: 4000-Hz carrier; blue: white-noise carrier; black: no stimulus). Error bars indicate 95% confidence intervals of the means.

A three-way repeated measures ANOVA [modulation frequency (2 levels) x carrier type (4 levels) x stimulus duration (6 levels)] with Geisser-Greenhouse corrected *p*-values was performed on the adjusted SNR. Modulation frequency, carrier type, and stimulus duration were treated as categorical within-subject factors. The analysis showed significant main effects of modulation frequency [$F(1, 14) = 143.74$, $p < 0.05$, $\eta_p^2 = 0.91$], carrier type [$F(3, 42) = 44.567$, $p < 0.05$, $\eta_p^2 = 0.76$], and stimulus duration [$F(5, 70) = 71.464$, $p < 0.05$, $\eta_p^2 = 0.84$] on the adjusted SNR. Significant interactions were also found between modulation frequency and carrier type [$F(3, 42) = 7.818$, $p < 0.05$, $\eta_p^2 = 0.36$], between modulation frequency and stimulus duration [$F(5, 70) = 19.304$, $p < 0.05$, $\eta_p^2 = 0.58$], between carrier type and stimulus duration [$F(15, 210) = 11.462$, $p < 0.05$, $\eta_p^2 = 0.45$], and between modulation frequency, carrier type, and stimulus duration [$F(15, 210) = 2.2444$, $p < 0.05$, $\eta_p^2 = 0.14$]. The effects of modulation frequency and stimulus duration on the adjusted SNRs were further explored by post hoc Dunn-Bonferroni analyses in different sessions for different experimental conditions.

Effects of carrier frequency

Post hoc comparisons showed that for 1000-Hz carrier, the adjusted SNRs for 40-Hz modulation frequency were significantly higher than those for 80-Hz modulation frequency for stimulus durations longer than 100 milliseconds ($p < 0.05$). For 40-Hz modulation frequency, the adjusted SNRs for 1000-Hz carrier were significantly larger than those of the control conditions once stimulus durations were longer than 100 milliseconds ($p < 0.05$). Also, the adjusted SNRs for 1000-Hz carrier increased significantly from 50 to 300 milliseconds in every step ($p < 0.05$); however after 300 milliseconds, the SNRs did not increase significantly every time stimulus duration was increased ($p > 0.05$). For 80-Hz modulation frequency, none of the adjusted SNRs for

1000-Hz carrier were significantly larger than those of the control conditions ($p > 0.05$). The adjusted SNRs for 1000-Hz carrier were not significantly different from that at 50 milliseconds until stimulus durations were longer than 300 milliseconds ($p < 0.05$).

Post hoc comparisons showed that for 4000-Hz carrier, the adjusted SNRs for 40-Hz modulation frequency were significantly higher than those for 80-Hz modulation frequency for stimulus durations of 200 and 400 milliseconds ($p < 0.05$). For 40-Hz modulation frequency, the adjusted SNRs for 4000-Hz carrier were significantly larger than those of the control conditions once stimulus durations were longer than 300 milliseconds ($p < 0.05$). Also, the adjusted SNRs for 4000-Hz carrier were not significantly larger than those at 50milliseconds until stimulus durations exceeded 200 milliseconds ($p < 0.05$); however after 200 milliseconds, the adjusted SNRs did not differ significantly until stimulus duration was increased to 800 milliseconds ($p < 0.05$). For 80-Hz modulation frequency, none of the adjusted SNRs for 4000-Hz carrier were significantly larger than those of the control conditions ($p > 0.05$). Also, none of the adjusted SNRs were significantly different between different stimulus durations for 4000-Hz carrier either ($p > 0.05$).

Pairwise comparisons showed that for white-noise carrier, the adjusted SNRs for 40-Hz modulation frequency were significantly higher than those for 80-Hz modulation frequency for stimulus durations of 200 and 400 milliseconds ($p < 0.05$). For 40-Hz modulation frequency, the adjusted SNRs for white-noise carrier were significantly larger than those of the control conditions once stimulus durations were longer than 300 milliseconds ($p < 0.05$). Also, the adjusted SNRs for white-noise carrier increased significantly from 50 to 200 milliseconds in every step ($p < 0.05$); however after 200 milliseconds, the SNRs did not increase significantly every time stimulus duration was increased ($p > 0.05$). For 80-Hz modulation frequency, none of the adjusted SNRs for

white-noise carrier were significantly larger than those of the control conditions ($p > 0.05$). Also, none of the adjusted SNRs were significantly different between different stimulus durations for white-noise carrier ($p > 0.05$). For both 40- and 80-Hz modulation frequencies, none of the adjusted response amplitudes were significantly different between different experimental conditions ($p > 0.05$).

In general, the adjusted SNRs for 40-Hz modulation frequency were larger than those for 80-Hz modulation frequency. For 40-Hz modulation frequency, the adjusted SNRs of the experimental conditions were larger than those of the control conditions once stimulus durations were longer than 300 milliseconds; however for 80-Hz modulation frequency, none of the adjusted SNRs of the experimental conditions were significantly different from those of the control conditions. For both 40- and 80-Hz modulation frequencies, there was no significant difference in the adjusted SNRs between different experimental conditions.

4.2.3. ASSR DETECTION

The percentage of responses detected was calculated for all stimulus conditions. A response was considered as present if its amplitude was larger than the mean amplitude of the adjacent noise components plus two standard deviations of the mean. The percentage of detected responses as a function of stimulus duration for all stimulus conditions is shown in Figure 12. For all experimental conditions, the percentage of detected responses increased as the stimulus duration increased after 200 milliseconds. Additionally, the percentages of detected responses were higher for 1000-Hz carrier and 40-Hz modulation frequency.

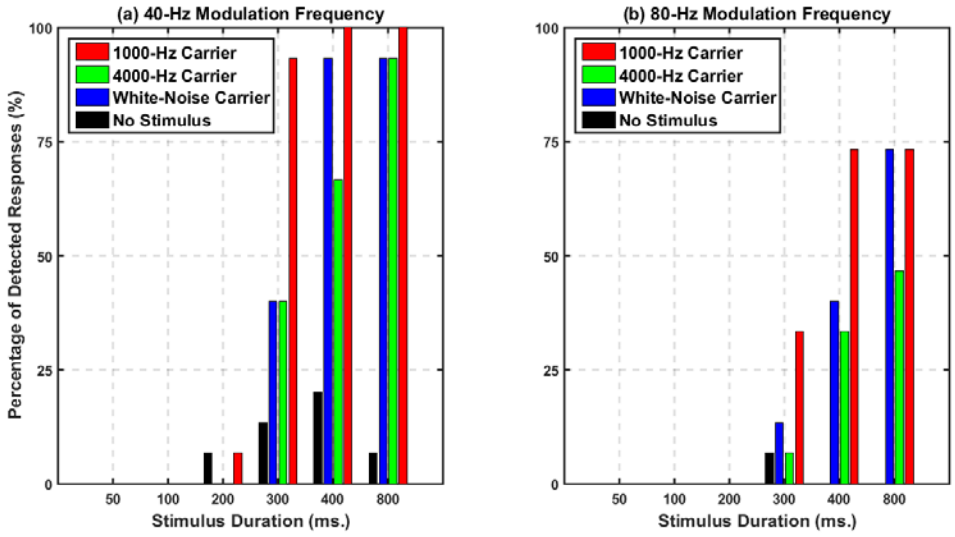

Figure 12 Percentage of detected responses plotted as a function of stimulus duration. Different panels represent different modulation frequencies (left: 40 Hz; right: 80 Hz). In each panel, different bar colors represent different carrier types (red: 1000-Hz carrier; green: 4000-Hz carrier; blue: white-noise carrier; black: no stimulus).

Based on the binomial distribution of the percentage of detected responses over stimulus duration, logistic regression curve $f(x) = \frac{1}{1+ e^{-(\beta_0 + \beta_1 x)}}$ (100%) was fitted to the data. Figure 13 shows the probability of detecting a response as a function of stimulus duration for both experimental and control conditions. The regression analysis showed that for both 40- and 80-Hz modulation frequencies, stimulus duration has a significant effect on the probability of detecting a response for the experimental conditions (p < 0.05), but not for the control conditions (p > 0.05). For 40-Hz modulation frequency, 1000-Hz carrier has a deeper curve slope (β_1 = 0.053) compared to those of 4000-Hz (β_1 = 0.012) and white-noise carriers (β_1 = 0.017). Additionally, the stimulus duration that corresponds to the 50% probability of detecting a response is 250 milliseconds for 1000-Hz carrier, 376 milliseconds for 4000-Hz carrier, and 330 milliseconds for white-noise carrier. For 80-Hz modulation frequency, curve slope (β_1) is 0.007 for 1000-Hz carrier,

0.005 for 4000-Hz carrier, and 0.007 for white-noise carrier, and the stimulus duration that corresponds to the 50% probability of detecting a response is 474 milliseconds for 1000-Hz carrier, 766 milliseconds for 4000-Hz carrier, and 589 milliseconds for white-noise carrier.

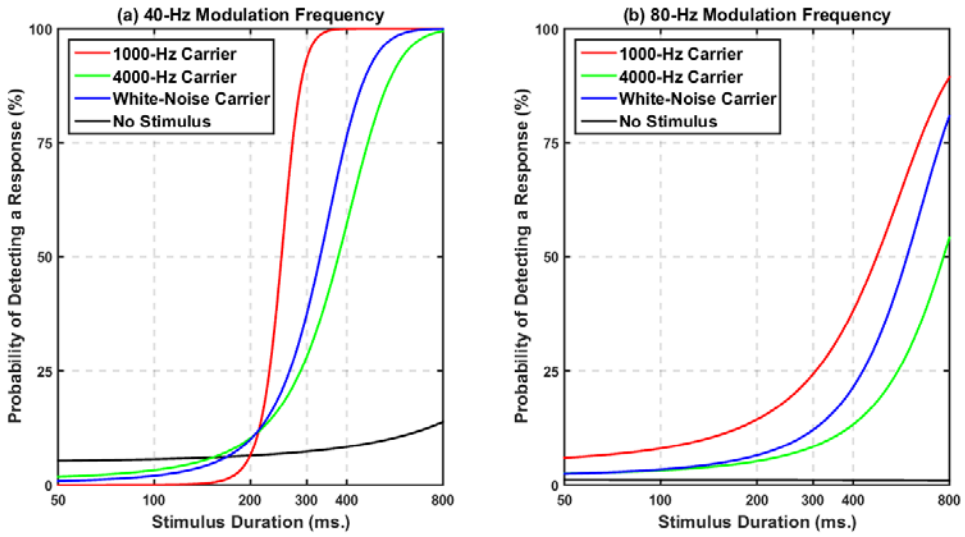

Figure 13 Probability of detecting a response plotted as a function of stimulus duration on a logarithmic scale. Different panels represent different modulation frequencies (left: 40 Hz; right: 80 Hz). In each panel, different line colors represent different carrier types (red: 1000-Hz carrier; green: 4000-Hz carrier; blue: white-noise carrier; black: no stimulus).

Maximum likelihood estimation was also used to fit a mixed logistic regression model for the percentage of detected responses, with fixed effects for modulation frequency and carrier type, and random intercept for subject. Modulation frequency and carrier type were added in sequence to a null model with only the random intercept for subject. Model comparisons through the ANOVA showed that the model fit was significantly improved by including modulation frequency ($\chi^2 = 23.938$, df = 1, $p < 0.05$) and carrier type ($\chi^2 = 79.62$, df = 3, $p < 0.05$) in the model. A cross-product term

(modulation frequency x carrier type) was next added to the model to test the possible interaction between modulation frequency and carrier type. The interaction was not statistically significant ($\chi^2 = 1.6212$, df = 3, $p > 0.05$).

With only the modulation frequency and carrier type in the model, the analysis showed that the percentages of detected responses were significantly higher for 40-Hz modulation frequency compared to 80-Hz modulation frequency for all experimental conditions ($p < 0.05$). For both 40- and 80-Hz modulation frequencies, the percentages of detected responses of all experimental conditions were significantly higher than those of the control conditions ($p < 0.05$). Additionally, among the experimental conditions, the percentages of detected responses were higher for 1000-Hz carrier than those for 4000-Hz and white-noise carriers ($p < 0.05$); however, there was no significant difference in the percentages of detected responses between 4000-Hz and white-noise carriers ($p > 0.05$).

4.2.4. ASSR PHASE CONTOUR

The instantaneous phases of both the ASSRs recorded under 800-ms stimulus conditions and their corresponding modulation signals were extracted through the Hilbert transform (HT) over time. The instantaneous phase of the ASSR was then subtracted from that of the modulation signal to derive a phase contour. The individual phase contours were adjusted so that the mean values between 300 and 700 milliseconds equal to zero degree. Figure 14 shows the mean adjusted phase plotted as a function of time with 95% confidence interval of the mean for the 800-ms stimulus conditions. In general, the confidence intervals of the phase contours were narrower for the experimental conditions compared to the control conditions. Compared to 80-Hz modulation frequency, the confidence intervals of the phase contours were narrower for 40-Hz

modulation frequency. For 40-Hz modulation frequency, the confidence intervals of the phase contours were narrower after 400 milliseconds, indicating stronger neural activity in response to the stimuli.

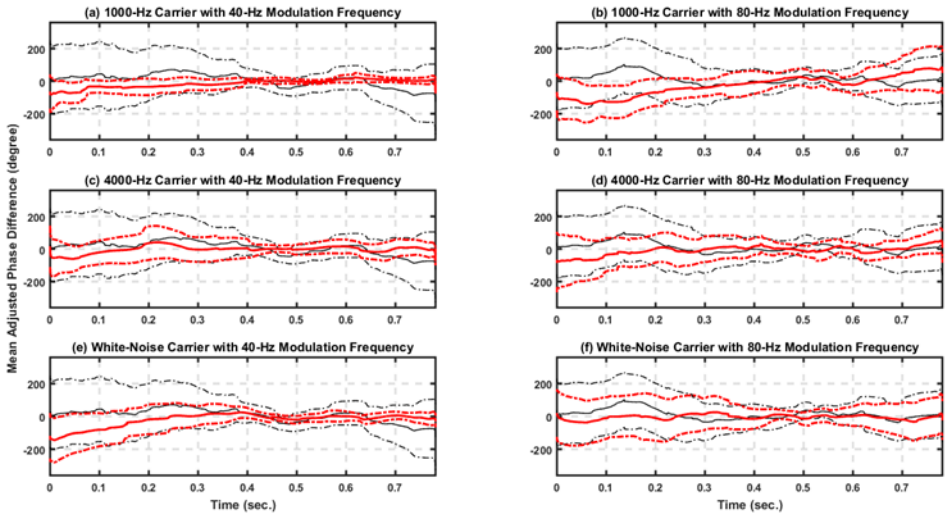

Figure 14 Mean adjusted phase difference between the 800-ms ASSR and modulation signal plotted as a function of time. Different columns represent different modulation frequencies (left: 40 Hz; right: 80 Hz), and different rows represent different carrier types (top: 1000-Hz carrier; middle: 4000-Hz carrier; bottom: white-noise carrier). In all panels, red line represents the mean adjusted phase contour of the experimental condition with red dashed lines indicating 95% confidence interval of the mean, and black line represents the mean adjusted phase contour of the control condition with black dashed lines indicating 95% confidence interval of the mean.

4.2.5. ASSR AMPLITUDE CONTOUR

The instantaneous amplitudes of the ASSRs recorded under the 800-ms stimulus conditions were also extracted through the HT over time. The individual 800-ms amplitude contours were then scaled so that the mean values between 300 and 700 milliseconds equal to the mean amplitude of all the amplitude contours in the same interval for the same stimulus condition. Figure 15 shows the mean adjusted

instantaneous amplitude of the 800-ms ASSR plotted as a function of time with 95% confidence interval of the mean for the 800-ms stimulus conditions. In general, the mean adjusted amplitude contours increased as the time passed for all the 800-ms stimulus conditions. The confidence intervals of the amplitude contours during the first 200 milliseconds of the stimulus duration were wider compared to the last 600 milliseconds. The decrease in variability with increasing duration indicates a more robust neural response. For modulation frequency of 40 Hz, the amplitude of the ASSR increased gradually over the first 200 milliseconds after the stimulus onset until reaching a steady-state plateau, and then dropped rapidly after the stimulus offset within 50 milliseconds. For modulation frequency of 80 Hz, the mean adjusted amplitude contours of the experimental conditions were not completely separated from that of the control conditions until the time passed 200 milliseconds. Compared to 40-Hz modulation frequency, the mean adjusted amplitude contours of the experimental conditions contained more fluctuations for 80-Hz modulation frequency.

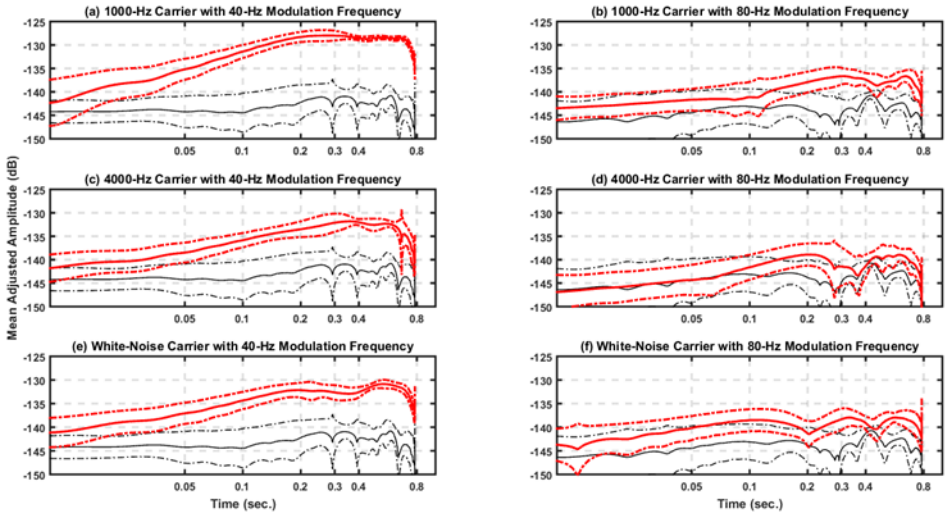

Figure 15 Mean adjusted instantaneous amplitude of the 800-ms ASSR plotted as a function of time on a logarithmic scale. Different columns represent different modulation frequencies (left: 40 Hz; right: 80 Hz), and different rows represent different carrier types (top: 1000-Hz carrier; middle: 4000-Hz carrier; bottom: white-noise carrier). In all panels, red line represents the mean amplitude contour of the experimental condition with red dashed lines indicating 95% confidence interval of the mean, and black line represents the mean amplitude contour of the control condition with black dashed lines indicating 95% confidence interval of the mean.

In order to mimic the trajectory of signal detection as a function of signal duration, the amplitude contours of the six 800-ms ASSRs were then reversed and shifted for each subject, so that individual mean amplitude contours between 300 and 700 milliseconds equal to zero dB. In other words, as the stimulus duration becomes longer, the ASSR amplitude becomes larger and the stimulus detection becomes easier. Three temporal integration models (linear, exponential, and power function models) for the experimental conditions were then evaluated by obtaining the best model fit of each to individual modified amplitude contours between 10 and 700 milliseconds. The initial parameter estimate of τ was set at 0.1 for both linear and exponential models, and the initial estimates of parameter values C and m were set at 0.9 and 0.8 for power function

model. For each 800-ms stimulus condition, a subject's modified amplitude contour was discarded if any of the three models failed to converge after 10,000 iterations. As a result, for 40-Hz modulation frequency, 14 subjects were included in the model fit for 1000-Hz carrier, and 15 subjects were included in the model fit for 4000-Hz and white-noise carriers. For 80-Hz modulation frequency, 13 subjects were included in the model fit for 1000- and 4000-Hz carriers, and 12 subjects were included in the model fit for white-noise carrier.

Table 1 shows means and standard deviations of the estimated parameter values and root mean square errors (RMSE) of the best-fitted least-squares functions from three temporal integration models for the 800-ms stimulus conditions. The means (standard deviations) of the Pearson product-moment correlations (r) between the best-fitted functions and modified amplitude contours were also provided. Based on the mean estimated parameter values, Figure 16 shows the best-fitted least-squares functions from three temporal integration models plotted as a function of time, along with the mean modified amplitude contour with 95% confidence interval of the mean for the 800-ms stimulus conditions. In general, three temporal integration models provided a better model fit (smaller RMSE values) to the modified amplitude contours for 40-Hz modulation frequency compared to 80-Hz modulation frequency. In particular for 40-Hz modulation frequency, exponential model seemed to provide a better fitting (smaller RMSE values) to the modified amplitude contours compared to both linear and power function models.

Table 1 Means (standard deviations) of the estimated parameter values and root mean square errors (RMSEs) of the best-fitted least-squares functions from three temporal integration models for the 800-ms stimulus conditions. The means (standard deviations) of the Pearson product-moment correlations (r) between the best-fitted functions and modified amplitude contours were also provided.

Modulation Frequency	**40 Hz**			**80 Hz**		
Carrier Type	**1000-Hz Tone**	**4000-Hz Tone**	**White Noise**	**1000-Hz Tone**	**4000-Hz Tone**	**White Noise**
Linear Model	*τ*: 0.106 (0.074) RMSE: 1.725 (0.950) r: 0.88 (0.11)	*τ*: 0.107 (0.116) RMSE: 2.361 (1.647) r: 0.66 (0.34)	*τ*: 0.082 (0.064) RMSE: 2.205 (0.808) r: 0.64 (0.35)	*τ*: 0.099 (0.066) RMSE: 3.063 (1.649) r: 0.50 (0.37)	*τ*: 0.105 (0.115) RMSE: 4.095 (1.920) r: 0.44 (0.34)	*τ*: 0.077 (0.100) RMSE: 2.692 (1.039) r: 0.42 (0.26)
Exponential Model	*τ*: 0.157 (0.085) RMSE: 1.405 (0.910) r: 0.89 (0.11)	*τ*: 0.161 (0.147) RMSE: 2.115 (1.760) r: 0.71 (0.29)	*τ*: 0.127 (0.090) RMSE: 2.000 (0.862) r: 0.70 (0.29)	*τ*: 0.154 (0.096) RMSE: 2.943 (1.520) r: 0.57 (0.29)	*τ*: 0.165 (0.148) RMSE: 3.962 (1.992) r: 0.52 (0.26)	*τ*: 0.120 (0.132) RMSE: 2.638 (1.043) r: 0.43 (0.27)
Power Function Model	*C*: 0.598 (0.161) *m*: 0.631 (0.339) RMSE: 1.525 (0.687) r: 0.81 (0.14)	*C*: 0.773 (0.551) *m*: 0.473 (0.343) RMSE: 2.160 (1.572) r: 0.60 (0.31)	*C*: 0.699 (0.250) *m*: 0.483 (0.336) RMSE: 1.982 (0.884) r: 0.62 (0.33)	*C*: 0.717 (0.228) *m*: 0.493 (0.314) RMSE: 2.791 (1.56r r: 0.53 (0.32)	*C*: 0.687 (0.246) *m*: 0.481 (0.333) RMSE: 3.818 (2.057) r: 0.44 (0.30)	*C*: 0.830 (0.178) *m*: 0.329 (0.286) RMSE: 2.637 (1.038) r: 0.41 (0.27)

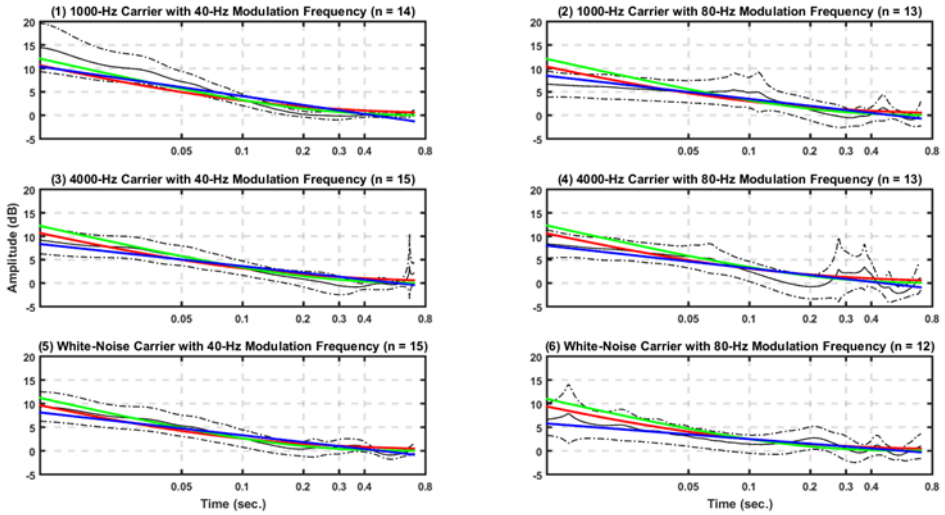

Figure 16 Best-fitted least-squares functions from three temporal integration models plotted as a function of time on a logarithmic scale. Different columns represent different modulation frequencies (left: 40 Hz; right: 80 Hz), and different rows represent different carrier types (top: 1000-Hz carrier; middle: 4000-Hz carrier; bottom: white-noise carrier). In each panel, different line colors represent predictions from different temporal integration models (red: linear model; green: exponential model; blue: power function model), with black line representing the mean modified amplitude contour of the 800-ms ASSR and black dashed lines indicating 95% confidence interval of the mean. In the title of each panel, the *n* value indicates the number of subjects included for model fitting.

V. DISCUSSION

5.1. ASSR compared to previous studies

In the current experiment, sinusoidally amplitude-modulated (SAM) signals with different carrier and modulation frequencies were used to evoke the auditory steady-state response (ASSR). For both 40- and 80-Hz modulation frequencies, the ASSR amplitudes for 1000-Hz carrier were larger than those for 4000-Hz carrier. Several studies have shown the ASSR amplitude decreases as the carrier frequency increases for both 40- and 80-Hz modulation frequencies (Galambos et al., 1981; John et al., 2001; John et al., 2002a; John et al., 2002b; Picton et al., 1987; Rodriquez et al., 1986; Roβ et al., 2003). Additionally, ASSR amplitudes in the current experiment were quite similar between 4000-Hz and white-noise carriers. However, the relationship between the ASSRs evoked by tone and noise carriers has not been previously reported in the literature.

Consistent with other studies, the ASSR amplitudes were greater for 40-Hz modulation frequency compared to 80-Hz modulation frequency. Previous studies (Aoyagi et al., 1999; Cohen et al., 1991; Rees et al., 1986) showed that the ASSR amplitude for 40-Hz modulation frequency was about two to three times (6-10 dB) larger than that for 80-Hz modulation frequency. In the current experiment, the ASSR amplitude for 40-Hz modulation frequency was more than twice (6.6 dB) the amplitude for 80-Hz modulation frequency. Moreover, the general findings for the ASSRs recorded in the present study are consistent with the results reported by other investigators (Aoyagi et al., 1999; Cohen et al., 1991; Galambos et al., 1981; John et al., 2001; John et al., 2002a; John et al., 2002b; Picton et al., 1987; Rees et al., 1986; Rodriquez et al., 1986; Roβ et al., 2003).

5.2. Potential Gamma-band artifact

Roß et al. (2002) showed that a transient gamma-band response (GBR) emerging in first 100 milliseconds after stimulus onset. The GBR is considered as one type of the long latency response (LLR) and can be recorded through the measurement of the auditory evoked potential (AEP) (Jacobson & Fitzgerald, 1997) and magnetoencephalography (MEG) (Pantev et al., 1991). The GBR typically consists of about 4 oscillations and usually emerges between 20 and 200 milliseconds following the stimulus onset (Jacobson & Fitzgerald, 1997). The energy of the GBR is primarily concentrated from 30 to 40 Hz (Pantev et al., 1991).

In the current experiment, no GBR was observed in the 800-ms ASSR. One of the possible reasons that no GBR was observed is because the inter-stimulus interval (ISI) is short at 43 milliseconds. This relatively short ISI allowed little for the auditory nervous system to recover before the next stimulus arrived. Several studies have reported that the amplitude of the GBR decreases when the ISI decreases (Pantev, 1995; Pantev et al, 1993; Roß et al., 2002). For example, Roß et al. (2002) found that the amplitude of the GBR was reduced by one third when the ISI was shortened from 3 to 0.5 seconds.

In order to confirm the GBR did not obscure the ASSR in 800-ms stimulus conditions in the current experiment, a followed-up experiment was conducted on five subjects. Four types of stimuli (i.e., 40-Hz SAM 1000-Hz tone, non-modulated 1000-Hz tone, 40-Hz SAM 4000-Hz tone, and non-modulated 4000-Hz tone) were delivered to each subject with the same experimental procedure used in the present study. The amplitude contour of the 800-ms ASSR under the condition with no modulation was subtracted from the one under the modulated condition for both 1000- and 4000-Hz carriers. No apparent morphological difference in the amplitude contours was observed

between difference wave and the original wave for both 40-Hz SAM 1000- and 4000-Hz tones. Therefore, it was concluded that the GBR was not a factor in the present study.

5.3. Duration effect

5.3.1. ASSR AMPLITUDE

In the current experiment, the ASSR amplitude was measured at the modulation frequency via the fast Fourier transform (FFT). The ASSR amplitude increased as the stimulus duration increased for both 40- and 80-Hz modulation frequencies. For 40-Hz modulation frequency, the amplitude of the ASSR began to increase at 50 milliseconds after the stimulus onset. In fact, the amplitude of the ASSR increased at an average rate of 5.5 dB for 1000-Hz carrier, 4.5 dB for 4000-Hz carrier, and 4.6 dB for white-noise carriers for every doubling of stimulus duration during the first 200 milliseconds, beyond which however, the amplitude of the ASSR increased only about 1 dB for every doubling of stimulus duration for all three carrier types. For 80-Hz modulation frequency, the amplitude of the ASSR did not begin to increase until about 200 milliseconds, beyond which the amplitude of the ASSR increased at an average rate of 3.3 dB for 1000-Hz carrier, 1.9 dB for 4000-Hz carrier, and 3.5 dB for white-noise carrier for every doubling of stimulus duration.

Several studies (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957) have shown behaviorally the detection threshold for tones decreased at an average rate of about 3 dB for every doubling of tone duration from about 10 to 250 milliseconds, after which the threshold change remained relatively small. Green et al. (1957) even showed that the rate was higher at 4.5 dB for signal durations shorter than 20 milliseconds and lower at 1.5 dB for signals longer than 100 milliseconds. For 40-Hz modulation frequency, the increasing rates of the ASSR amplitude over the first 200 milliseconds and

the overall trends of the ASSR amplitude development after 200 milliseconds were consistent with previous studies (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957).

5.3.2. ASSR SIGNAL-TO-NOISE RATIO

The amplitude data were also considered relative to the noise that was simultaneously present by examining the signal-to-noise ratio (SNR). The SNR of the ASSR increased as the stimulus duration increased for both 40- and 80-Hz modulation frequencies. For 40-Hz modulation frequency, the SNR of the ASSR began to increase at 50 milliseconds after the stimulus onset. In fact, the SNR of the ASSR increased at an average rate of 3.7 dB for 1000-Hz carrier, 2.9 dB for 4000-Hz carrier, and 3.0 dB for white-noise carrier for every doubling of stimulus duration. For 80-Hz modulation frequency, the SNR of the ASSR increased little during the first 200-ms interval, beyond which the SNR of the ASSR increased at an average rate of 4.0 dB for 1000-Hz carrier, 3.0 dB for 4000-Hz carrier, and 3.8 dB for white-noise carrier for every doubling of stimulus duration.

The increasing rates of the ASSR signal-to-noise ratio over the first 200 milliseconds were also consistent with previous studies (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957); however, no horizontal segment was revealed after 200 milliseconds. In fact, the overall trend of the ASSR signal-to-noise ratio development was more consistent with the study by Florentine et al. (1988). In the Florentine et al. (1988) study, no horizontal segment of the behavioral integration function as predicted by Plomp and Bouman, (1959) was found. Although the ASSR amplitude and SNR for 80-Hz modulation frequency did not increase until about 200 milliseconds due to weaker

responses overall, the increasing rates beyond 200 milliseconds were about the same as those for 40-Hz modulation frequency.

5.3.3. ASSR DETECTION

In the present study a response was considered as present if its amplitude exceeded the mean amplitude of the adjacent noise components plus two standard deviations. In general, the percentages of ASSRs detected increased rapidly as the stimulus duration increased beyond 200 milliseconds for both 40- and 80-Hz modulation frequencies. This criterion used to judge the presence of a neural response seems to be a great way to determine whether the amplitude of a neural response has grown to its full potential. However, compared to the gradual behavioral detection threshold improvement over the first 250 ms of stimulus duration (Hughes, 1946; Plomp & Bouman, 1959; Green et al., 1957), if the development of the ASSR over time for 40-Hz modulation frequency correlates with the behavioral temporal integration, it is less likely that the human auditory system uses the same criterion as a method for signal detection.

5.3.4. AMPLITUDE CONTOURS

In the current experiment, the instantaneous amplitudes of the ASSRs were extracted through the Hilbert transform (HT). The overall shape of the adjusted amplitude contour, measured in the condition of 1000-Hz carrier with 40-Hz modulation frequency, is similar to the one recorded in the study by Roß et al. (2002). The amplitude of the ASSR increased gradually over the first 200 milliseconds until reaching the steady state, and then dropped rapidly after the stimulus offset within 50 milliseconds. Additionally, the confidence interval of the amplitude contour during the first 200 milliseconds of the stimulus duration was wider compared to the last 600 milliseconds. The decrease in variability with increasing duration indicates a more robust neural response. In fact,

similar patterns were also observed on the adjusted amplitude contours for 4000-Hz and white-noise carriers with 40-Hz modulation frequency. For 80-Hz modulation frequency, the ASSR amplitude also increased as the time increased for all three experimental conditions. However, the adjusted amplitude contours were less distinguishable from those in the control condition due to weaker responses overall.

5.3.5. PHASE CONTOURS

In the present study, the phenomenon of temporal integration was investigated by comparing the phases of the ASSRs to those of the corresponding modulation signal of the eliciting stimuli over time. It is thought that if a neural response was present, the ASSR would be phase-locked to the stimulus, and the phase difference between the ASSR and the modulation signal would remain relatively constant over time. If no response was present, the phase would vary randomly. For both 40- and 80-Hz modulation frequencies, the relatively narrow confidence interval for the phase contour indicates the likelihood of a response compared to the control condition. For 40-Hz modulation frequency, the phase of the ASSR was strongest when the stimulus reached the duration of 400 milliseconds, beyond which the phase remained relatively stable. On the other hand, for 80-Hz modulation frequency the phase locking was less stable across subjects, which suggests weaker responses. In the context of response phase, the phenomenon of temporal integration was more clearly revealed in the ASSRs obtained in the condition of 40-Hz modulation frequency compared to those recorded in the condition of 80-Hz modulation frequency in terms of neural phase-locking.

In general, temporal integration could be observed on the amplitude and phase development of the ASSR over the first 200 milliseconds after the stimulus onset. These results were consistent with previous results from both psychoacoustic (Hughes, 1946;

Plomp & Bouman, 1959; Green et al., 1957) and physiologic (Roß et al., 2002) studies. The first 200 milliseconds of the amplitude contour may represent the process of temporal integration. The decrease in variability with increasing duration indicates a more robust neural response. However, once the stimulus was terminated, the auditory system returned back to its original state in a relatively short amount of time. The last few milliseconds of the amplitude contour might be considered as the recovery time of the auditory system. Only the signals that fall within the interval of the first 200 milliseconds after the stimulus onset would be integrated together to improve the detectability of the signal. On the other hand, only the signals that appear after the auditory system has been completely recovered would be treated as different stimuli. This point of view can be used to explain why the auditory system, which is considered as a leaky integrator, has both the long and short time constants as mentioned in the beginning of the article. On the other hand, the confidence interval of the phase contour decreased dramatically after 400 milliseconds compared to the first 400 milliseconds, indicating that phase-locked neural activities became stronger once the stimulus duration was longer than 400 milliseconds.

5.4. Models of temporal integration

Three temporal integration models (i.e., linear, exponential, and power function models) were used to fit the modified amplitude contours for ASSRs acquired with 40- and 80-Hz modulation frequencies. These models were chosen because they were clearly defined and widely used in the past to examine temporal integration data measured behaviorally. The exponential model was observed to provide a slightly better fit (smaller root mean square error, RMSE) to the modified amplitude contours compared to the linear and power models. Among the three carriers, white-noise carrier had a slightly

shorter time constant (τ) and lower exponent (m) compared to those for 1000- and 4000-Hz carriers, indicating that white-noise carrier had the fastest temporal integration time.

In the present study, the average time constant estimated by the exponential model (Plomp & Bouman , 1959) was approximately 0.160 seconds for both 1000- and 4000-Hz carriers. This time constant was lower than the one reported by Plomp and Bouman (1959) but higher than the one calculated by Watson and Gengel (1969). More specifically, Plomp and Bouman (1959) showed that the time constant was about 0.33 seconds for 1000-Hz tone and 0.23 seconds for 4000-Hz tone. On the other hand, Watson and Gengel (1969) reported that the time constant was about 0.08 seconds for 1000-Hz tone and 0.05 milliseconds for 4000-Hz tone. However in the current experiment, no apparent frequency effect on the time constant was observed across three temporal integration models, which was consistent with the findings from Florentine et al. (1988). In fact, the time constants for 4000-Hz carrier were slightly larger than those for 1000-Hz carrier. It is speculated that because the ASSR evoked by the SAM 4000-Hz tone is typically smaller in amplitude compared to the ASSR evoked by the SAM 1000-Hz tone, the phenomenon of temporal integration is more easily observed for 1000-Hz carrier than that for 4000-Hz carrier and therefore no obvious frequency effect on temporal integration was revealed in the ASSR.

5.5. Temporal integration

In the present study, the phenomenon of temporal integration was observed in the ASSR for both 40- and 80-Hz modulation frequencies. The ASSR amplitude for 40-Hz modulation frequency increased steadily over the first 200 milliseconds after the stimulus onset, beyond which the amplitude remained relatively constant; however for 80-Hz modulation frequency, the ASSR amplitude did not begin to increase until about 200

milliseconds due to weaker responses overall. Additionally, for both 40- and 80-Hz modulation frequencies the narrower confidence interval for the phase contour indicates the likelihood of a response compared to the control condition, especially after 400 milliseconds. Because previous studies showed that the neural generators of the ASSR are primarily located in the auditory midbrain (Hari et al., 1989; Herdman et al., 2002; Johnson et al., 1988; Kiren et al., 1994; Mäkelä et al., 1990; Spydell et al., 1985) and brainstem (Cone-Wesson et al., 2002; Hari et al., 1989; Herdman et al., 2002; Kiren et al., 1994; Mäkelä et al., 1990; Picton et al., 2003) when the modulation frequency is 40 Hz and 80-Hz respectively, the results from the present study suggested that temporal integration could be observed in the ASSR from both neural cites in the auditory nervous system.

In the present study, the phenomenon of temporal integration can be observed in the ASSR from listeners with normal hearing for both 40- and 80-Hz modulation frequencies. However, it is not clear whether the listener is making of the cues revealed in the ASSR amplitude and phase, or operating on the SNR development over time for making decisions about signal detection. Additionally, it is not clear whether a listener with a fast temporal integration rate behaviorally would have a comparably fast temporal integration rate electrophysiologically. Therefore, future studies will focus on the correlation between electrophysiological and behavioral test results, discovering the exact strategy that listeners use for signal detection in the auditory system.

Additional to the typical population, the phenomenon of temporal integration can be investigated with an atypical population through the measurement of the ASSR. One example of an atypical population is patients with auditory-related learning problems. Patients with auditory-related learning problems have been clinically diagnosed with learning disability or attention deficit disorder but exhibit poor performance on auditory

processing battery tests even though their hearing sensitivity (threshold < 20 dB HL for octaves frequencies from 500 to 4000 Hz) were normal. Many of these participants were found to have auditory temporal and spectral processing abnormalities (Wright et al., 1997). In addition, it has been demonstrated that the representation of the fundamental frequency information in the brainstem response is impaired in children with autism spectrum disorders (Russo et al., 2008; Russo et al., 2009). These patients are usually difficult to test with conventional behavioral auditory tests. Therefore, investigating the phenomenon of temporal integration on these populations through the measurement of the ASSR might reveal findings different from those in the typical population.

5.6. Conclusions

The results from the present study suggested that the phenomenon of temporal integration could be observed in the ASSR. Evidence of temporal integration was more clearly revealed when the ASSR was recorded under 40-Hz modulation frequency compared to 80-Hz modulation frequency; however, no apparent frequency effect on the time constant was revealed. Among three temporal integration models, exponential model seemed to fit the electrophysiological data best. Future studies will focus on the correlation between behavioral and electrophysiological test results and the differences in the ASSRs obtained from typical and atypical populations.

References

Alain, C., Woods, D., & Covarrubias, D. (1997). Activation of duration-sensitive auditory cortical fields in humans. Electroencephalography and Clinical Neurophysiology, 104, 531-539.

Aoyagi, M., Suzuki, Y., Yokota, M., Furuse, H., Watanabe, T., & Ito, T. (1999). Reliability of 80-Hz amplitude-modulation-following response detected by phase coherence. *Audiology and Neurootology*, 4, 28-37.

Arnold, S. A. (2000). The auditory brain stem response. In: Roeser, R. J., Valente, M., & Hosford-Dunn, H (Ed.), *Audiology: Diagnosis* (pp. 451-470). New York, NY: Thieme.

Basar, E., Basar-Eroglu, C., & Greitschus, F. (1987). The association between 40 Hz-EEG and the middle latency response of the auditory evoked potential. *International Journal of Neuroscience*, 33, 103-117.

Bohórquez. J, & Ozdamar, O. (2008). Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. *Clinical Neurophysiology*, 119, 2598-2607.

Buus, S. (1999). Temporal Integration and multiple looks, revisited: weights as a function of time. *Journal of the Acoustical Society of America*, 105, 2466-2475.

Champlin, C. A. (1992). Methods for detecting auditory steady-state potentials recorded from humans. *Hearing Research*, 58, 63-69.

Cohen, L. T., Richards, F. W., & Clark, G. M. (1991). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *Journal of the Acoustical Society of America*, 90, 2467-2479.

Cone-Wesson, B., Dowell, R. C., Tomlin, D., Rance, G., & Ming, W. J. (2002). The auditory steady-state response: comparisons with auditory brainstem response. *Journal of the Acoustical Society of America*, 13, 173-187.

Dimitrijevic, A., & Roβ, B. (2008). Neural generators of the auditory steady-state response. In G. Rance (Ed.), *Auditory Steady-State Response: Generation, Recording, and Clinical Applications*. (pp. 83-108). San Diego, CA: Plural.

de Boer , E. (1985). Auditory time constants: A paradox? In: Michelsen, A. (Ed.), *Time Resolution in Auditory Systems.* (pp. 141-158), Berlin: Springer.

Florentine, M., Fastl, H., & Buus, S. (1988). Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. *Journal of the Acoustical Society of America*, 84, 195-203.

Gage, N., & Roberts, T. (2000). Temporal integration: reflections in the M100 of the auditory evoked field. *NeuroReport*, 11, 2723-2726.

Galambos, R., Makeig, S., & Talmachoff, P. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences*, USA 78, 2643-2647.

Garner, W.R. (1947). The effect of frequency spectrum on temporal integration in the ear. *Journal of the Acoustical Society of America*, 19, 805-815.

Geisler, C. D. (1960). Average responses to clicks in man recorded by scalp electrodes. *Technical Report 380*. Cambridge, MA: Massachusetts Institute of Technology Research Laboratory of Electronics.

Gerken, G., Bhat, V., & Hutchison-Clutter, M. (1990). Auditory temporal integration and the power function model. *Journal of the Acoustical Society of America*, 88, 767-778.

Green, D. M. (1960). Auditory detection of a noise signal. *Journal of the Acoustical Society of America*, 32, 121-131.

Green, D. M. (1985). Temporal factors in psychoacoustics. In: Michelsen, A. (Ed.), *Time Resolution in Auditory Systems* (pp. 122-140). Berlin: Springer-Verlag.

Green, D. M., Birdsall, T., & Tanner, W. (1957). Signal detection as a function of signal intensity and duration. *Journal of the Acoustical Society of America*, 29, 523-531.

Hari, R., Hämaläinen, M., Joutsíniemi, S. L. (1989). Neuromagnetic steady-state responses to auditory stimuli. *Journal of the Acoustical Society of America*, 86, 1033-1039.

Herdman, A. T., Lins, O. S., van Roon P, Stapells, D. R., Scherg, M., & Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, 15, 69-86.

Hood, L. J. (1998). *Clinical Applications of the Auditory Brainstem Response*. San Diego, CA: Singular.

Hughes, J. (1946). The thresholds of audition for short periods of stimulation. *Proceedings of the Royal Society*, London B133, 486-490.

Jacobson, G. P., & Fitzgerald, M. B. (1997). The auditory evoked gamma band potential (aeGBP) in normal subjects. *Journal of the Acoustical Society of America*, 8, 44-52.

Jerger, J., Chmiel, R., Frost, J. D., & Coker, N. (1986). Effect of sleep on the auditory steady state evoked potential. *Ear and Hearing*, 7, 240-245.

John, M. S., Dimitrijevic, A., & Picton, T. W. (2002a). Auditory steady-state responses to exponential modulation envelopes. *Ear and Hearing*, 23, 106-117.

John, M. S., Dimitrijevic, A., van Roon, P., & Picton, T. W. (2001). Multiple auditory steady-state responses to AM and FM stimuli. *Audiology and Neurootology*, 6, 12-27.

John, M. S., & Picton, T. W. (2000). MASTER: A Windows program for recording multiple auditory steady-state responses. *Computer Methods and Programs in Biomedicine*, 61, 125-150.

John, M. S., Purcell, D. W., Dimitrijevic, A., & Picton, T. W. (2002b). Advantages and caveats when recording steady state responses to multiple simultaneous stimuli. *Journal of the American Academy of Audiology*, 13, 246-259.

Johnson, B. W., Weinberg, H., Ribary, U., Cheyne, D. O., & Ancill, R. (1988). Topographic distribution of the 40 Hz auditory evoked-related potential in normal and aged subjects. *Brain Topography*, 1, 117-121.

Keidel, W. D., Kallert, S., & Korth, M. (1983). *The Physiological Basis of Hearing*. New York, NY: Thieme-Stratton.

Kiren, T., Aoyagi, M., Furuse, H., & Koike, Y. (1994). An experimental study on the generator of amplitude-modulation following response. *Acta Oto-laryngologica Supplement*, 511, 28-33.

Levi, E. C., Folsom, R. C., & Dobie, R. A. (1993). Amplitude-modulation following response (AMFR): effects of modulation rate, carrier frequency, age, and state. *Hearing Research*, 68, 42-52.

Lins, O. G., & Picton, T. W. (1995). Auditory steady-state response to multiple simultaneous stimuli. *Electroencephalography and Clinical Neurophysiology*, 96, 420-432.

Loveless, N., & Hari, R. (1993). Auditory evoked fields covary with perceptual grouping. *Biological Psychology*, 35, 1-15.

Loveless, N., Hari, R., Hämäläinen, M., & Tiihonen, J. (1989). Evoked responses of human auditory cortex may be enhanced by preceding stimuli. *Electroencephalography and Clinical Neurophysiology*, 74, 217-227.

Loveless, N., Levänen, S., Jousmäki, V., Sams, M., & Hari, R. (1996). Temporal integration in auditory sensory memory: neuromagnetic evidence. *Electroencephalography and Clinical Neurophysiology*, 100, 220-228.

Mäkelä, J. P., Karmos, G., Molnar, M., Csépe, V. & Winkler, I. (1990). Steady-state responses from the cat auditory cortex. *Hearing Research*, 45, 41-50.

Onishi, S., & Davis, H. (1968). Effect of duration and rise time of tone-bursts on evoked V-potentials. *Journal of the Acoustical Society of America*, 44, 582-591.

Pantev, C. (1995). Evoked and induced gamma-band activity of the human cortex. *Brain Topography*, 76:321:330.

Pantev, C., Elbert, T., Makeig, S., Hampson, S., Eulitz, C., & Hoke, M. (1993). The relationship of transient and steady-state auditory evoked fields. *Electroencephalography and Clinical Neurophysiology*, 90, 82-90.

Pantev, C., Eulitz, C., Elbert, T., & Hoke, M. (1994). The auditory evoked sustained field: origin and frequency dependence. *Electroencephalography and Clinical Neurophysiology*, 90:82-90.

Pantev, C., Makeig, S., Hoke, M., Galambos, R., Gallen, C., & Hampson, S. (1991). Human auditory evoked gamma band magnetic fields. *Proceedings of the National Academy of Sciences*, USA 88, 8896-9000.

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state response. *International Journal of Audiology*, 42, 177-219.

Picton, T. W., Skinner, C. R., Champagne, S. C., Kellett, A. J., & Maiste, A. C. (1987). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *Journal of the Acoustical Society of America*, 82, 165-178.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978). Human auditory sustained potentials. I. The nature of the response. *Electroencephalography and Clinical Neurophysiology*, 45, 186-197.

Plomp, R., & Bouman, A. (1959). Relation between hearing threshold and duration for tone pulses. *Journal of the Acoustical Society of America*, 31, 749-758.

Plourde, G., Stapells, D. R., Picton, T.W. (1991). The human auditory steady-state evoked potentials. *Acta Oto-laryngologica Supplement*, 491, 153-159.

Rees, A., Green, G. G., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research*, 23, 123-133.

Regan, D. (1989). *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York, NY: Elsevier.

Roβ, B., Draganova, R., Picton, T. W., & Pantev, C. (2003). Frequency specificity of 40-Hz auditory steady-state responses. *Hearing Research*, 186, 57-68.

Roβ, B., Picton, T. W., & Pantev, C. (2002). Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hearing Research*, 165, 68-84.

Rodriguez, R., Picton, T., Linden, D., Hamel, G., & Laframboise, G. (1986). Human auditory steady state responses: Effects of intensity and frequency. *Ear and Hearing*, 7, 300-313.

Russo, N. M., Nicol, T., Trommer, B., Zecker, S., & Kraus, N. (2009). Brainstem transcription of speech is disrupted in children with autism spectrum disorders. *Developmental Science*, 12, 557-567.

Russo, N. M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., & Kraus, N. (2008). Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology*, 119, 1720-1731.

Sahley, T. L., Nodar, R. H., & Musiek, F. E. (1997). *Efferent auditory system*. San Diego, CA: Singular.

Santarelli, R., & Conti, G. (1999). Generation of auditory steady-state responses: linearity assessment. *Scandinavian Audiology*, 28, 23-32.

Sheft, S., & Yost, W. A. (1990). Temporal integration in amplitude modulation detection. *Journal of the Acoustical Society of America*, 88, 796-805.

Spydell, J. D., Pattee, G., & Goldie, W. D. (1985). The 40 Hz auditory event-related potential: normal values and effects of lesions. *Electroencephalography and Clinical Neurophysiology*, 62, 193-202.

Stapells, D. R., Linden, D., Suffield, J. B., Hamel, G., & Picton, T. W. (1984). Human auditory steady state potentials. *Ear and Hearing*, 5, 105-113.

Sussman, E., Winkler, I., Ritter, W., Alho, K., & Näätänen, R. (1999). Temporal integration of auditory stimulus deviance as reflected by the mismatch negativity. *Neurosicence Letters*, 264, 161-164.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. Journal *of the Acoustical Society of America*, 66, 1364-1380.

Viemeister, N. F., & Wakefield, G. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, 90, 858-865.

Warr, W. B. (1980). Efferent components of the auditory system. *Annals of Otorhinolaryngology, 89 (Suppl.)*, 114-120.

Warr, W. B. (1992). Organization of olivocochlear efferent systems in mammals. In: Webster, D. B., Popper, A. N., & Fay, R. R. (Ed.), *The Mammalian Auditory Pathway: Neuroanatomy* (pp. 410-448). New York, NY: Springer-Verlag.

Watson, C. S., & Gengel, R. W. (1969). Signal duration and signal frequency in relation to auditory sensitivity. *Journal of the Acoustical Society of America*, 46, 989-997.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S. Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, 387, 176-178.

Zwislocki, J. (1960). Theory of temporal auditory summation. *Journal of the Acoustical Society of America*, 32, 1046-1060.

Zwicker, E., & Wright, H. N. (1963). Temporal summation for tones in narrow-band noise. *Journal of the Acoustical Society of America*, 35, 691-699.

Vita

Yu-Fu Chen was born in Taipei, Taiwan and graduated from Taipei Chenggong High School in 2003. After receiving the Bachelor of Science in Speech Language Pathology and Audiology from Chung Shang Medical University in 2007, he performed his military service and served as an educational audiologist in National Taichung School for Students with Hearing Impairments. Between 2009 and 2016, he went overseas and attended the University of Texas at Austin, obtaining both Doctor of Audiology and Doctor of Philosophy in Communication Sciences and Disorders after intensive clinical and research trainings. During these years, he was awarded the Jesse H. Jones Endowed Centennial Fellowship for his strong interdisciplinary ability.

Email: audifisher@gmail.com

This dissertation was typed by the author.